MODELING OF NON-FULLY DEPLETED SILICON-ON-INSULATOR MOSFETS, AND APPLICATIONS TO HIGH-PERFORMANCE/LOW-POWER ULSI DESIGN

BY

DONGWOOK SUH

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1995

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude and appreciation toward my advisor, Professor Jerry G. Fossum, for his devoted guidance, patient encouragement, and unceasing support for my academic and personal affairs throughout my graduate study. By his invaluable guidance, this work was made possible. I also would like to thank the members of my supervisory committee, Professors R. M. Fox, M. E. Law, S. S. Li, and C. C. Hsu for their willing service and guidance on my committee.

I am grateful to Texas Instrument Inc., IBM, and the Semiconductor Research Corporation for their technical and financial support throughout my research. Especially, I feel obliged to express my deep thanks to Dr. Ted Houston at TI and Mr. Mario Pelella at IBM for providing technological information invaluable to the integrity of this work.

I would like to recognize in this achievement Messrs. Ping-Chin Yeh and Srinath Krishnan, who helped me in many ways with the sincere and profound discussions on the related topics. Gratitude is also extended to many of my friends, who helped me through technical discussions or by making my life one of the most cheerful. I cannot mention all, but I would like to mention Drs. Jin-Young Choi, Joohyun Jin, Haeseok Cho, Ghy-Boong Hong, Keith R. Green,

Taehoon Kim, and Messrs. Dukhyun Chang, Kwang-Rip Hyun, Jaewan Kim, and Timothy H. Lim.

I am very pleased that my lovely wife and critic, JinJoo and my dearest daughter Josephine Jung-Ah participated in this achievement with their unyielding support. Without them, the long years of my graduate study would have been hard and barren. Last, but not the least, I would like to share this work with my father, Dr. Nam Won Suh. His examples in life and academics from my early ages made me plan this pursuit. I am very happy that now I am able to hang my diploma under his.

TABLE OF CONTENTS

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

MODELING OF NON-FULLY DEPLETED SILICON-ON-INSULATOR MOSFETS, AND APPLICATIONS TO HIGH-PERFORMANCE/LOW-POWER ULSI DESIGN

By

Dongwook Suh

August 1995

Chairman: Professor Jerry G. Fossum
Major Department: Electrical and Computer Engineering

This dissertation presents a physical yet compact model for non-fully depleted (NFD) silicon-on-insulator (SOI) MOSFETs. The high-low junction in the body of a typical NFD/SOI MOSFET mimics the charge condition of that in thin-film accumulated (TFA) model formalism. Therefore, we base our new model on the TFA formalism and then expand it to reflect the device characteristics pertinent to deep submicron short-channel NFD/SOI MOSFETs. The channel-current analysis is done for all inversion conditions. The model also includes the generation current due to the non-local impact ionization and the thermal generation, and the parasitic-BJT and recombination currents in the device which are important in the short-channel devices, especially when operating in the floating-body mode. LDD/LDS and halo analyses are done separately and later linked to the channel

analysis by a numerical method to completely characterize the device's current-voltage relations. These analyses further define the terminal charges for modeling the charging currents in the device. The analyses are characterized with model parameters which directly relate to the device structure and physics. The resulting model is implemented in the framework of SPICE2G.6 to create an enhanced version of SOISPICE. SOISPICE model predictions are then verified using device measurement data and numerical simulation results. The new NFD/SOI MOSFET model in SOISPICE is used to examine the effect of body configuration. With simple breakdown conditions derived from SOISPICE simulations, it is shown that the body resistance of body-tied devices must be impractically low for efficacy. The floating-body configuration, however, is subject to many floating-body effects, including transient ones. Two floating-body effects--the thermal generation-driven threshold lowering and the transient suppression of off-state leakage--can be beneficial in circuit performance. However, the instabilities due to the dynamic threshold variation and the transient BJT current require design conservatism in certain circuit applications, as exemplified by SOISPICE simulations. With the insight from SOISPICE simulations, an NFD/SOI DRAM is assessed, concentrating on the dynamic floating-body charge; the viability of SOI technology in DRAM application is suggested.

CHAPTER ONE
INTRODUCTION

Since the first fabrication of integrated circuits (ICs) based on bulk-silicon complementary-metal-oxide-semiconductor (CMOS) technology, the scaling down of the feature device size has been pursued for better performance and more areal efficiency. While the economy of the scaling has been proved historically [Bre80], the further miniaturization and its anticipated benefits commensurate with the past are in doubt when the device is scaled into the deep-submicron region [Now93]. The challenges in deep-submicron scaling are now twofold: the control of the so-called short-channel effect (SCE) and the enhancement of IC performance, now including the total on-chip power consumption in addition to speed. The problems of the SCE are almost, if not completely, solved by the local introduction of high doped regions in the device [Cod85, Nag92, Hor92], but performance still portends a greater challenge. The main dilemma is the tradeoff between on-chip power consumption and speed of the IC system. High-performance/low-power devices, which are generally desirable for any IC systems, are not readily achieved via bulk-silicon device technology. For example, higher doping levels in the device for SCE control tend to increase the power consumption due to increased parasitic capacitance. Also, the reduction of the supply voltage for lowering

switching power and enhancing reliability sacrifices performance since the device threshold voltages can not be reduced arbitrarily due to stand-by power limitations.

Silicon-on-insulator (SOI) technology has been contending the established bulk-silicon technology in this regard. Experimentally as well as theoretically [Sha92, Fos93a], MOSFET devices built on SOI substrates show superior performance, mainly due to the reduced source/drain parasitic capacitances and the easier control of the short-channel effects. The reduction of the parasitic capacitances is estimated up to 30% [Hu93], and the junction depth, which plays a critical role in controlling SCE in bulk-silicon technology, is implicitly limited by the thin SOI layer. Furthermore, the immunity of the SOI devices to single-event upset (SEU), which is very critical in ULSI circuits, enables more aggressive scaling in several applications under radiative environment. In addition to the aforementioned device merits, the inherent device isolation afforded by the underlying back oxide contributes to a simpler manufacturing process and more dense integration by eliminating the need for wells in the CMOS technology.

In establishing SOI CMOS technology, there have been two major trends, i.e., fully depleted (FD) and non-fully depleted (NFD), or partially depleted (PD) SOI MOSFETs. While both modes of device operation take on the inherent advantages over the bulk-silicon MOSFET, FD devices were favored in the early stage of SOI development because of their performance advantages including enhanced current drivability and ideal subthreshold

slope. These advantages, over NFD/SOI and bulk-silicon MOSFETs, tend to fade, however, as the channel length is scaled; the current enhancement is limited by velocity saturation [Fos93b] and the subthreshold slope is degraded due to source/drain charge sharing and increasing significance of the surface states [Yeh93]. In addition to these fundamental limitations, acceptable threshold voltage, which is to be decided by an essential performance/power tradeoff consideration [Yeh94], dictates higher body doping for the FD technology, which is also needed to control the short-channel effects. Then to maintain the full-depletion condition, the SOI film must be scaled down accordingly, with the acceptable thickness for the deep-submicron feature length projected below 50 nm [Yeh94]. However, this approach not only requires a tighter control over the material preparation, but also brings up increased parasitics such as source/drain and contact resistance [Hwa93]. All of these facts cast real doubts on further scalability of the FD technology.

Consequently, there is a growing interest in the NFD SOI technology in which the SOI film is thick enough to ensure non-full depletion. Because of the thicker SOI film, NFD design allows vertical doping variation [Yan93] and/or halo implantation (DI-LDD) [Cod83], which have been incorporated in scaled bulk-silicon technology for the control of the short-channel effect, and the threshold voltage is easily controlled by the channel doping. In addition, an NFD device is more immune to radiation than an FD device because of the neutral body that separates the channel from the back Si-SiO_2 interface. However, a drawback exists: due to the neutral portion of the body

that floats, and the sensitivity of threshold voltage to the floating-body voltage, the floating-body effects are more prevalent in NFD devices and circuits than in FD or bulk counterparts. The most obvious of them is a saturation kink [Col88], which is an increased saturation current/conductance caused by impact ionization. This effect is known to be detrimental in analog applications [Sha93], while purportedly beneficial to digital applications as argued in [Sha93]. Thus the design in the NFD technology must involve investigating these new effects, which were nonexistent or of little consequence in the bulk technology.

For design in NFD SOI CMOS, there is a tendency to use bulk-silicon MOSFET models, empirically modified to apply to the SOI device. The reliability of such models however is questionable due to the empiricism, and thus they do not provide physical insights that can be very useful in technology development. For example, capacitance associated with the underlying SiO_2 can be important if the oxide is scaled, and floating-body effects due to (DC or transient) free-carrier charging can influence, or even control [Sha93], the device performance. Indeed, the floating-body effects can be detrimental or beneficial, and hence must be inhibited or should be exploited in optimal design of the NFD technology. However, the device performance merits do not translate into circuit performance enhancements in a straightforward manner, and hence the assessment of the floating-body effects must be done comprehensively at the circuit as well as the device levels.

A physical, yet compact model for the NFD/SOI MOSFET is thus needed to account for the unique physical nature of the device. This dissertation presents such a model and uses it, implemented in SOISPICE, to give insights regarding optimal device/circuit design for the NFD technology.

In Chapter 2, the physical and compact charge-based model is developed through modifications and extensions of existing SOI and bulk MOSFET models [Vee88], [Cho91a], and [Gre93a]. The model is applicable down to the deep-submicron channel length (L ~ 0.1μm), including the short-channel effects such as drain-induced barrier lowering and current enhancement (DIBL and DICE), charge sharing, and channel-length modulation.

The model development is facilitated by using a regional analyses (i.e., in extrinsic and intrinsic regions), the linking of which constitutes the complete formalism. The acronym, NFD, is derived from the fact that the SOI film thickness is greater than the maximum depletion width induced by the MOSFET gate, and therefore the silicon layer is never fully depleted. However, the model extension from existing TFA (thin-film-accumulated) formalism [Vee88], which assumes full depletion of SOI layer, is possible because of the high-low junction in NFD devices that mimic the back surface charge and bias condition in TFA. The strong-inversion model in TFA is then extended to include weak- and moderate-inversion characterization to complete the channel current model. The parasitic-BJT and the impact-ionization currents are included, which manifest themselves in short-channel device characteristics. Carrier generation and recombination are important mecha-

nisms in the NFD device model, as they characterize the unique floating-body effects associated with the terminal charge dynamics. The developed model is modular and hierarchical; the various components in the model and the structural variations of the device can be effected at the user's discretion to give further insights regarding device/circuit design.

In Chapter 3, a detailed description of the implementation of the model in SOISPICE is provided, followed by the model verification. The implementation details include the network representation of the model, the model algorithms, and the description of the functions of the newly added SPICE model routines. The verification is based on numerical device simulator results as well as measured data.

The verification accompanies the model parameter-extraction scheme which obviates the necessity of rigorous optimization, even though the measured data are influenced by self-heating. The model predictions, with parameters based on the device structure and a few measured characteristics, are compared to 2-D device simulation results and measurement data from differently scaled technologies. They show generally a very good match, including the floating-body effects, but with noticeable discrepancies in self-heating-affected portions of the measured data.

In Chapter 4, SOISPICE is used to analyze the efficacy of body ties. The body ties are introduced in NFD/SOI technology with an intention to suppress the floating-body effects revealed in Chapter 3. However, the efficacy of them is in question as is apparent from common body-tied device characteris-

tics such as the premature-BJT breakdown. The effectiveness is discussed with the breakdown voltages in device characteristics as a criterion; the conditions are derived with simplifying assumptions rendered by our NFD/SOI device model. These conditions are then interpreted in terms of the device parameters to give insights regarding the body-tied device design and the NFD/SOI technology as well.

In Chapter 5, the floating-body effects in steady and transient circumstances are analyzed in NFD/SOI devices as well as in circuits. As revealed in Chapter 3, the floating body causes numerous so-called floating-body effects in device characteristics in steady state even below breakdown, some being purportedly beneficial. We deal with the questions of the inhibition and/or exploitation of these effects in a viable NFD/SOI CMOS technology, using device/circuit simulations with SOISPICE, and noting relevant parametric sensitivities implied by the simulations. Indeed, the steady-state floating-body effects do not translate into circuit-level merits or demerits in a straightforward manner as they are compounded in transient situations. In that sense, the unique accountings of the floating-body charge dynamics effected in SOISPICE with our new model, and the simulations using it, present and solve these problems, as the instabilities in circuit operations due to transient floating-body effects are revealed in the latter part of the chapter.

In Chapter 6, with the insights about floating-body effects gained from Chapter 5, we address the issues in SOI DRAM design. A successful DRAM design on an SOI substrate could advance the full-blossomed emer-

gence of SOI technology because of its prodigious production volume. However, in spite of many technological and electrical advantages over bulk DRAMs which drew early attention, DRAMs on SOI also show unstable dynamic data retention. We begin by providing insights concerning DRAM cell transistor design by characterizing the dynamic retention time in the floating-body mode. Then we address a generalized methodology for floating-body SOI DRAM design, which is used to project the viability of the technology.

In Chapter 7, this dissertation is concluded with a summary and suggestions for future works to enhance the NFD/SOI model.

CHAPTER TWO
A PHYSICAL MODEL FOR THE NON-FULLY DEPLETED (NFD) SOI MOSFET

2.1 Introduction

This chapter presents a physical model for the NFD SOI MOSFET. The model is developed through modifications and extensions of the existing SOI and bulk MOSFET models [Vee88, Cho91a, Gre93a]. The model developed here is charge-based, and targeted for the deep-submicron device, including short-channel effects such as the threshold voltage variation due to charge sharing, drain-induced current enhancement (DICE), drain-induced barrier lowering (DIBL), channel-length modulation, and impact ionization. It also accounts for the parasitic bipolar transistor (BJT) in the SOI MOSFET structure.

Figure 2.1 shows a representation of the typical NFD SOI MOSFET with halo implant. The intrinsic body is composed of two regions with different doping levels (P^- and P), forming a high-low junction. The N regions represent LDS and LDD. The P^+ regions around the LDS and LDD are the halo. The highly doped halo is included to prevent punchthrough and to control the threshold voltage, but it also influences other device characteristics dependent on body doping. Also shown in Fig. 2.1 is a P^+ strip at the end of the source region. It represents one method to realize the

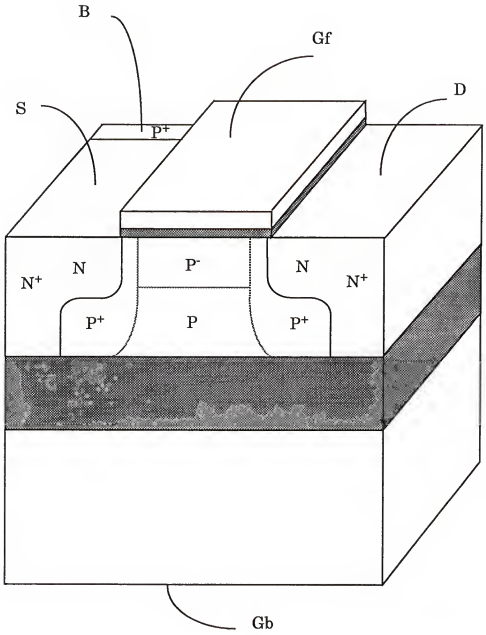


Fig. 2.1 The structure of a non-fully depleted (NFD) SOI MOSFET. LDD/LDS regions are denoted as N. The P^+ regions under the LDD and LDS represent the halo regions. P^+ region in the source at the back of the device represents the body contact which may exist in symmetry to form a double body-tied-to-source (DBTS) structure. Without any contact to the body, the device will operate in the floating-body mode.

intrinsic body-source tie, whose spacing and structure define a finite body resistance, R_B. When P^+ strips do not exist, the body is left electrically floating, arousing numerous floating-body effects subject to the internal source-body voltage. Since R_B is typically high, floating-body effects can be seen even when the body is tied.

The SOI film is thick enough in the NFD device to ensure that it is not fully depleted at strong inversion. However, we base our model on the TFA formalism [Vee88] for thin-film devices because the charge condition at the high-low junction mimics that of TFA devices. The TFA model is, however, insufficient to characterize the complete physical nature of the NFD/halo device in Fig. 2.1, and must be expanded. To completely characterize the physical nature of the NFD device with halo implant, the model extension and modification must include physical accountings for subthreshold current, effects of LDS and LDD and impact-ionization current, as well as the charges that define charging currents. Parasitic bipolar effects are modeled as well because of the possibility of dynamic BJT effects as well as the breakdown and latch mechanisms associated with the BJT. The halo influences these effects. Proper modeling for thermal-generation/recombination currents must also be stressed in the new device model because of their importance in predicting the floating-body effects.

To effect the most physical characterization, properly linked regional analyses constitute the model. The intrinsic channel, and the LDS and LDD regions are characterized separately; the LDD analysis is linked to

the channel analysis iteratively. The parasitic BJT and the impact ionization influence this characterization. The current-voltage solutions from the analyses provide information to characterize the terminal charges.

Capable of simulating the variations of the structure in Fig. 2.1, the model is hierarchical and modular; the electrical and the structural effects can be optionally included or excluded to give users the advantages of being able to investigate a variety of structures as indicated in Fig. 2.2. The entire modeling is done for ultimate implementation into SOISPICE [Fit91], which will be discussed in the next chapter.

2.2 Model Formalism

Figure 2.3 shows a schematic cross-section of a typical (n-channel) NFD/SOI MOSFET with lightly doped drain/source (LDD/LDS) and halo implant, which is the basis of our model derivation. It is a five-terminal device when the body is contacted. Floating body is a model option, as are the LDD/LDS and halo regions. Unlike the case in Fig. 2.1, the doping density in each region of Fig. 2.3 is assumed uniform. The intrinsic body is assumed to comprise two regions, defined by the vertical doping variation and separated by a high-low junction as indicated. The low-doped region (N_{BL}) is assumed to be fully depleted when the channel is formed while the high-doped region (N_{BH}) is assumed to remain neutral. These model assumptions pertaining to the doping are representative of the body structure that results from vertical

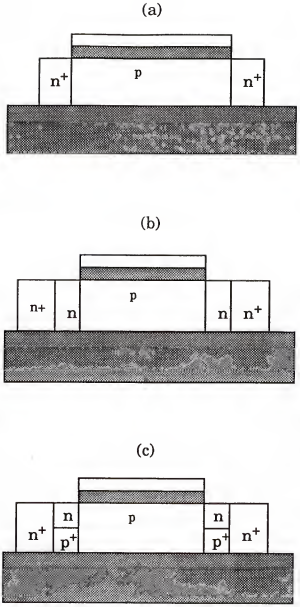


Fig. 2.2 Three hierarchical structures incorporated in the model.
(a) without LDD/LDS and halo;
(b) without halo but with LDD/LDS;
(c) with halo and LDD/LDS

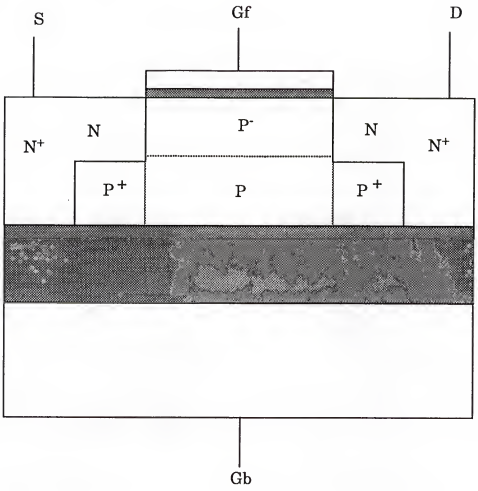


Fig. 2.3 The base structure of model derivation. The doping densities in the figure are assumed uniform in each region.

doping engineering [Yan92] for control of the short-channel effect and prevention of back-channel turn-on.

The high-low junction in the NFD body effectively confines the depletion region under the front gate, and the underlying quasi-neutral region decouples the back gate electrically. Thus the electrical properties of the channel region in the NFD device mimic the TFA (thin-film-accumulated) condition [Vee88] in which accumulated holes at the back interface prevent the back gate from electrically interfering with the channel. Figure 2.4 illustrates the similarities in electrical properties of the NFD and TFA devices. If the high-low junction were ideal (i.e., infinite N_{BH}), the charge condition in the NFD device would be identical to that of a TFA device, with the back surface potential (Ψ_{sb}) pinned at the body-source voltage V_{BS}. Consistent with the model assumptions, we define an effective film thickness (t_b), which is (approximately) the high-low junction depth or the maximum depletion-region width in strong inversion, and use the TFA formalism [Vee88] as the basis for the NFD model. Having accounted for the nonideality of the high-low junction by the definition of t_b, we now assume in our model that Ψ_{sb}, the body potential with respect to the neutral ground [Lim83], is equal to V_{BS}.

The TFA short-channel-current (I_{CH}) analysis is expanded to account for weak- and moderate-inversion conditions. The intrinsic channel is analyzed for strong and weak inversion according to the bias conditions, and the moderate-inversion mode of operation is effected by linking both regions

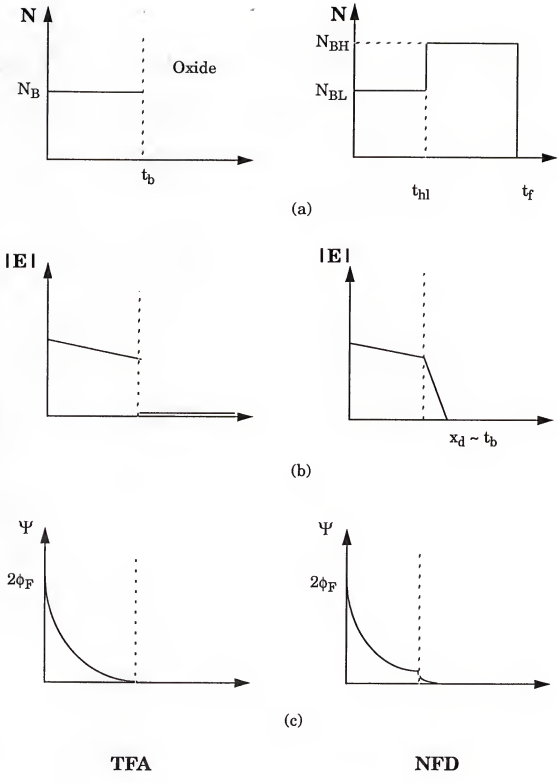


Fig. 2.4 The comparison between TFA and NFD.
(a) Doping profile; (b) Electric Field; (c) Potential

via cubic-spline interpolation. To completely characterize the physical properties of the NFD/SOI MOSFET, impact-ionization current (I_{Gi}), parasitic-BJT current (I_T), recombination current (I_R), and thermal-generation current (I_{Gt}) are also included in the model, and terminal charging currents are defined to simulate the charge dynamics. The generation/recombination currents and the body charging current are crucial for proper accounting of the floating-body effects on channel current as well as parasitic-BJT current, and must be symmetric with respect to the source and drain.

To facilitate the model development, we separately characterize the intrinsic region under the gate and the optional extrinsic regions (LDD/LDS, halo), linking the regional solutions numerically. The accounting for the halo region is done by defining an effective (higher) N_{BL} which yields a higher threshold voltage. The LDD/LDS analyses define the voltage drops across the LDD and LDS. The ohmic portion of these drops is simply modeled by a lumped resistance added to the parasitic series resistance in the source and drain, while a possible non-ohmic component for the LDD, due to high electric field and concomitant depletion, is accounted for physically. This voltage drop is subject to the analysis of the channel. Thus an iterative solution scheme is adopted. The solution defines quasi-static terminal charges, as well as the mentioned currents, which are used to characterize the charging/discharging currents.

2.3 Channel Current

The channel current is derived from the characterization of the charges under the front gate. The derivation can be facilitated by considering two distinctive conditions in which either drift or diffusion dominates the conduction mechanism, i.e. by assuming strong or weak inversion. The moderate-inversion characteristics are represented by connecting these two regions via cubic-spline interpolation [Bur81]. The boundaries of the moderate-inversion region are physically defined by the dependence of the channel charge, Q_c, on the front-gate bias, V_{GfS}, following Tsividis [Tsi82]. This dependence is derived by applying Gauss's law to the front-gate structure for $V_{DS} = 0$ [Vee88]:

$$Q_{C0} = -C_{of}\left[V_{GfS} - V_{FB}^{f} - (1+\alpha)\,\Psi_{sf0} + \frac{Q_{b(eff)}}{2C_{of}} + \alpha V_{BS}\right] \qquad (2.1)$$

where $\alpha = C_b/C_{of}$ when $C_b = \varepsilon_s/t_b$ and $C_{of} = \varepsilon_{ox}/t_{oxf}$, and Ψ_{sf0} is the surface potential. V_{FB}^{f} is the front-gate flat-band voltage and $Q_{b(eff)}$ is the effective (areal) body charge density. To correctly characterize (2.1) for all inversion cases and thus to define the physical moderate-inversion boundaries, the short-channel effects such as charge sharing and the drain-induced current enhancement and barrier lowering (DICE and DIBL) must be incorporated, and they are described first.

2.3.1 Charge Sharing

Charge-sharing analysis is meaningful only in strong inversion, and is modeled with the uniform effective doping density in the low-doped body. It is modeled by defining triangles of charges as in Fig. 2.4 which represent the depletion charge controlled by source and drain. The base of the triangle, d, is approximated by the depletion width of the abrupt pn junction. The potential drop across the junction is equal to V_{bi}-V_{BS} where V_{bi} is the built-in potential of S/D-B pn junction, calculated using N_{BL} and N_{LDS} (doping density in LDD/LDS) if the LDD/LDS are part of the device. Now applying pn junction theory [Sze81], we get

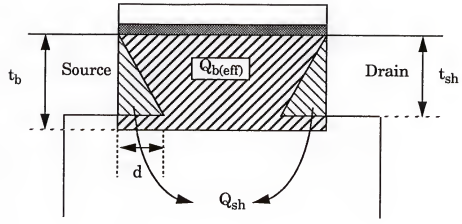
$$d = \frac{V_{bi}^{s,d} - V_{BS}}{E_{b(eff)}} \frac{N_{LDS}}{N_{LDS} + N_{BL}} \tag{2.2}$$

where

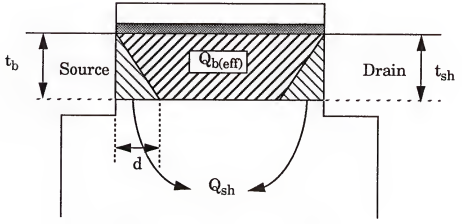
$$E_{b(eff)} = \sqrt{\frac{qN_{BL}\left(V_{bi}^{s,d} - V_{BS}\right)}{2\varepsilon_s} \frac{N_{LDS}}{N_{LDS} + N_{BL}}}. \tag{2.3}$$

If LDS/LDD are not part of the device, the above analysis is done with N_{LDS} replaced by N_{DS} (doping density in the high-doped S/D).

Now the amount of the charge controlled by the source is $Q_{sh} = -qN_{BL}Wt_{sh}d/2$ and the same amount is controlled by the drain for symmetry. The value of t_{sh} may vary according to the device structure. The NFD device can have the LDS thickness smaller or larger than t_b. Thus we



(a)



(b)

Fig. 2.5 Schematic diagram of a simple charge-sharing model.
(a) $t_b > t_{LDS}$; (b) $t_b < t_{LDS}$.

define t_{sh} to be equal to t_{LDS} when t_b is larger than t_{LDS}, or to t_b otherwise, as shown in Fig. 2.5.

Now subtracting the depletion charge controlled by source and drain from $Q_b = -qN_{BL}t_b$, we obtain the effective body depletion charge density,

$$Q_{b\,(eff)} = Q_b - \frac{Q_{sh}}{WL} = -qN_B t_b \left(1 - \frac{dt_{sh}}{Lt_b} \right). \tag{2.4}$$

2.3.2 DIBL and DICE

In short-channel MOSFETs, the applied drain voltage, V_{DS}, affects the potential and electric field in the channel and body such that modulation of channel charge and/or perturbation of the surface potential occurs. In weak inversion, V_{DS} directly modulates the surface potential, thereby increasing the channel current. In strong inversion, V_{DS} not only affects the surface potential but directly gives rise to more channel charge which results in current enhancement. The former effect is called DIBL and the latter DICE. Consequently the channel current cannot be modeled by the gradual-channel approximation (GCA) alone.

To model DIBL and DICE, we must solve the 2-D Laplace's equation for the incremental change of the potential in the depleted body region [Vee88]:

$$\frac{\partial^2}{\partial x^2}\Delta\Psi + \frac{\partial^2}{\partial y^2}\Delta\Psi = 0\,. \tag{2.5}$$

Assuming a quasi separation of variables, we approximate (2.5) as

$$\frac{\partial^2}{\partial x^2}\Delta\Psi = -\frac{\partial^2}{\partial y^2}\Delta\Psi = -\eta \tag{2.6}$$

where $\eta = (2/L^2)V_{DS}$ since $\Delta\Psi(0,L)=V_{DS}$. Integrating (2.6) twice along x, we get

$$\Delta\Psi_{sb}(y) = \Delta\Psi_{sf}(y) - \Delta E_{sf}(y)\, t_b - \frac{\eta t_b^2}{2} \tag{2.7}$$

In NFD devices, Ψ_{sb} ($=V_{BS}$) is not directly affected by V_{DS}. Thus $\Delta\Psi_{sb} \cong 0$, and (2.7) can be rearranged as

$$\Delta E_{sf}(y) = \frac{\Delta\Psi_{sf}(y)}{t_b} - \frac{\eta t_b}{2} \tag{2.8}$$

Applying Gauss's law to the front surface using (2.8), assuming a vertical field in the region, we define the incremental charge-potential relation

$$\Delta Q_C = C_{of}(1+\alpha)\Delta\Psi_{sf} - \frac{\varepsilon_s t_b \eta}{2} \tag{2.9}$$

In (2.9), the second term of the right-hand side represents the channel charge enhancement in strong inversion due to drain bias which is not included in the GCA model. This effect is called DICE. In weak inversion, the variation of the channel charge due to the drain bias is negligible compared to the variation of potential, i.e., $\Delta Q_C \cong 0$, and (2.9) is modified to represent the DIBL effect:

$$\Delta\Psi_{sf} = \frac{\varepsilon_s t_b \eta}{2C_{of}(1+\alpha)}. \tag{2.10}$$

Note that the variation of the surface potential in DIBL is independent of y, consistent with the assumption of diffusion-dominant current in weak

inversion.

2.3.3 Moderate-Inversion Boundaries

The boundaries of the moderate-inversion region are physically defined by the dependence of the channel charge, Q_C, on the front-gate bias V_{GfS}, following Tsividis [Tsi82]. By rearranging (2.1), we relate the surface potential to the gate bias for $V_{DS} = 0$:

$$V_{GfS} = V_{FB} + (1+\alpha)\Psi_{sf0} - \frac{Q_{b(eff)}}{2C_{of}} - \frac{Q_{C0}}{C_{of}} - \alpha V_{BS}. \tag{2.11}$$

Differentiating both sides of (2.11) with respect to V_{GfS}, we get

$$\frac{d\Psi_{sf}}{dV_{GfS}} = \frac{C_{of}}{C_{of} + C_b + C_i} \tag{2.12}$$

where $C_i = d|Q_C|/d\Psi_{sf}$. Similar to the criteria of Tsividis [Tsi82], the lower limit of strong inversion is determined by the constancy of $d|Q_C|/dV_{GfS}$ and the upper limit of weak inversion by the constancy of the slope, $d(\ln|Q_C|)/dV_{GfS}$. With (2.12), we can express these derivatives as

$$\frac{d|Q_C|}{dV_{GfS}} = \frac{C_{of}C_i}{C_{of} + C_b + C_i} \tag{2.13}$$

and

$$\frac{d(\ln|Q_C|)}{dV_{GfS}} = \frac{C_{of}}{C_{of} + C_b + C_i}\frac{C_i}{|Q_C|}. \tag{2.14}$$

In (2.13) and (2.14), C_i is the only capacitance that varies with bias. Thus we define the moderate-inversion boundaries such that $C_i = 10(C_{of}+C_b)$

in (2.13) and $C_i = (C_{of}+C_b)/100$ in (2.14). These conditions can be translated to terms of surface potential and gate bias if we express C_i as [Sze81]

$$C_i = \frac{d|Q_C|}{d\Psi_{sf}} = \frac{qn_i^2}{N_{BL}}\frac{1}{E_{xf0}}\exp\left(\frac{\Psi_{sf}}{V_{TH}}\right) \tag{2.15}$$

where V_{TH} is the thermal voltage (kT/q) and E_{xf0} is the transverse electric field at the surface. In weak inversion, the average field is given simply by the depletion approximation as

$$E_{xf0} \cong \frac{\Psi_{sf0}-V_{BS}}{t_b} - \frac{Q_b}{2\varepsilon_s}. \tag{2.16}$$

In strong inversion, the channel charge affects the field, which is obtained from Poisson's equation,

$$\frac{d^2\Psi_{sf}}{dx^2} = \frac{q}{\varepsilon_s}\left[N_{BL} + \frac{n_i^2}{N_{BL}}\exp\left(\frac{\Psi_{sf}}{V_{TH}}\right)\right], \tag{2.17}$$

Integrating (2.17) over the predominant inversion layer, from $\Psi=\Psi_{sf0}$ to $\Psi=2\phi_F$, describes the surface field as

$$E_{xf0}^2 \cong E_x^2(2\phi_F) + \frac{2qn_i^2V_{TH}}{\varepsilon_s N_{BL}}\exp\left(\frac{\Psi_{sf}}{V_{TH}}\right) \tag{2.18}$$

where $E_X(2\phi_F) \cong (2\phi_F - V_{BS})/t_b - Q_{b\,(eff)}/2\varepsilon_s$ and $\phi_F = V_{TH}\ln(N_{BL}/n_i)$.

Substituting (2.16) and (2.18) into (2.14) and (2.13) respectively, with aforementioned criteria, we obtain the (V_{DS} = 0) moderate-inversion boundaries as

$$\Psi_{sfW} = V_{TH} \ln\left[\frac{N_{BL}E_{xf0}}{100qn_i^2}(C_{of} + C_b)\right] \tag{2.19}$$

and

$$\Psi_{sfS} = V_{TH} \ln\left[\frac{10N_{BL}E_{xf0}}{qn_i^2}(C_{of} + C_b)\right] \tag{2.20}$$

where Ψ_{sfW} is the upper limit of weak inversion and Ψ_{sfS} is the lower limit of strong inversion. Note that E_{xf0} in (2.16) and (2.18) is dependent on Ψ_{sf}; therefore a few iterations are required to calculate Ψ_{sfS} and Ψ_{sfW} in (2.19) and (2.20).

Due to DIBL and DICE in short-channel devices, the surface potential and the channel charge are modulated by the drain bias, and thus the boundaries that separate strong- and weak- inversion can also depend on V_{DS}. While DIBL directly increases the surface potential in weak inversion by $\varepsilon_s \eta t_b / 2C_{of}(1+\alpha)$, DICE gives rise to increased channel charge under the gate in strong inversion. These effects result in rigid shifts of the values of V_{GfS} that correspond to (2.19) and (2.10). These values are defined from (2.9), (2.10), (2.11), (2.19), and (2.20):

$$V_{TW} = V_{FB} + (1+\alpha)\Psi_{sfW} - \frac{Q_b}{2C_{of}} - \alpha V_{BS} - \frac{\varepsilon_s \eta t_b}{2C_{of}} \tag{2.21}$$

and

$$V_{TS} = V_{FB} + (1+\alpha)\Psi_{sfS} - \frac{Q_{b(eff)}}{2C_{of}} - \frac{Q_{C0}}{C_{of}} - \alpha V_{BS} - \frac{\varepsilon_s \eta t_b}{2C_{of}} \tag{2.22}$$

Q_C is neglected in (2.21) because of weak inversion, and in (2.22) it is calcu-

lated by Gauss's law using (2.17). Note that $Q_{b(eff)}$ in (2.22) is less than Q_b due to charge sharing in strong inversion.

2.3.4 Strong-Inversion Current

When V_{GfS} is greater than V_{TS}, the channel current is calculated by the strong-inversion model. For strong inversion, the carrier transport in the channel is predominantly due to the carrier drift and thus can be expressed as

$$I_{CH} = Wv|Q_C| \tag{2.23}$$

where v is carrier velocity in the channel and Q_C is the total aerial channel charge density. The total charge density Q_C is equal to $Q_{C0}+\Delta Q_C$, where Q_{C0} is given by (2.1) with Ψ_{sf0} pinned at Ψ_{sfS}, and ΔQ_C is given by (2.9).

2.3.4.1 Carrier Velocity-Field Model

Due to possible high longitudinal electric field as well as the high transverse field, the carrier velocity-field characteristics cannot be modeled by a simple linear relation. For a sufficiently high longitudinal electric field, the velocity tends to saturate. Thus, following [Vee88], we use a piecewise continuous model for the carrier velocity in the channel:

$$\begin{aligned} v(y) &= \frac{\mu_{eff}E_y}{1+\frac{\mu_{eff}E_y}{2v_{sat}}} && v(y) \le v_{sat} \\ &= v_{sat} && \text{otherwise,} \end{aligned} \tag{2.24}$$

where the longitudinal field is defined as $E_y = |d\Psi/dy|$. In (2.24), μ_{eff}, the low-

(longitudinal)-field mobility, is dependent on the transverse field in the channel. This dependence is modeled [Gar87] by the average of the transverse field as

$$\mu_{eff} = \frac{\mu_{n0}}{1 + \theta \overline{E}_X(y)}, \tag{2.25}$$

where θ is a fitting constant, and μ_{n0} is the zero-field mobility. The average transverse field is defined as $\overline{E}_{xf} = \overline{E}_{xf0} + \Delta\overline{E}_{xf}$ where $\overline{E}_{xf0}$ is the field when $V_{DS}=0$ and $\Delta\overline{E}_{xf}$ is the variation in the field due to $V_{DS}>0$. $\overline{E}_{xf0}$ is calculated as

$$\overline{E}_{xf0} = \frac{2\phi_F - V_{BS}}{t_b} - \frac{Q_{b(eff)}}{2\varepsilon_s} - \frac{Q_{C0}}{2\varepsilon_s} \tag{2.26}$$

and $\Delta\overline{E}_{Xf}$ is computed as

$$\Delta\overline{E}_{xf} = \frac{C_{of}\Delta\Psi_{sf}}{2\varepsilon_s}(\alpha - 1) - \frac{\eta t_b}{4}. \tag{2.27}$$

By rearranging (2.25), we can express μ_{eff} as

$$\mu_{eff} = \frac{\mu}{1 - B\Delta\Psi_{sf}} \tag{2.28}$$

where

$$\mu = \frac{\mu_{n0}}{1 + \frac{\theta C_{of}}{2\varepsilon_s}\left[-\frac{Q_{C0}}{C_{of}} - \frac{Q_{b(eff)}}{C_{of}} + \frac{2C_b}{C_{of}}(\Psi_{sfS} - V_{BS}) - \frac{C_b}{C_{of}}\left(\frac{t_b}{L}\right)^2 V_{DS}\right]} \tag{2.29}$$

and

$$B = \frac{\theta C_{of}}{2\varepsilon_s}\frac{\mu}{\mu_{n0}}(1-\alpha)\,. \tag{2.30}$$

Note that μ and B are now bias-dependent but spatially constant.

2.3.4.2 Triode Region Current

In the triode region, v(L) is smaller than v_{sat} and using (2.24), we derive the steady-state channel current as

$$I_{CH} = -WvQ_C = -WQ_C(y)\frac{\mu_{eff}\dfrac{\partial \Psi_{sf}}{\partial y}}{1+\dfrac{\mu_{eff}}{2v_{sat}}\dfrac{\partial \Psi_{sf}}{\partial y}}\,. \tag{2.31}$$

Following [Vee88], we rearrange (2.31) to obtain

$$I_{CH} = \frac{W\mu_{eff}\left[Q_C^2(0) - Q_C^2(L)\right]}{2LC_{of}(1+\alpha)\left(1+\dfrac{\mu_{eff}V_{DS}}{2v_{sat}L}\right)} \tag{2.32}$$

where μ_{eff} is defined as

$$\mu_{eff} = \frac{\mu}{1-f_B B V_{DS}} \tag{2.33}$$

with f_B being an empirical parameter ($0<f_B<0.5$) that reflects the nonlinearity of $\Delta\Psi_{sf}(y)$.

2.3.4.3 Saturation Current

As V_{DS} increases, a high longitudinal electric field occurs near the drain and causes the carrier velocity to saturate. Then the channel is divided into two regions where the carrier velocity is saturated and where the velocity

is field-dependent. At the boundary between two regions, $y = L_e$ ($=L-y_D$), we define $V_{DS(eff)} = \Delta\Psi_s(L_e)$ ($<V_{DS}$). The saturation current is then characterized by equating the currents in both regions as in [Vee88]. Equating (2.32), with L and V_{DS} replaced by L_e and $V_{DS(eff)}$, and (2.23) with $v = v_{sat}$, we obtain the saturation current description:

$$I_{CH} = \frac{W\mu_{eff}\left[Q_C^2(0) - Q_C^2(L_e)\right]}{2L_eC_{of}(1+\alpha)\left(1+\frac{\mu_{eff}V_{DS(eff)}}{2v_{sat}L_e}\right)}$$

$$= -Wv_{sat}Q_C(L_e). \tag{2.34}$$

From (2.34) we get

$$V_{DS(eff)} = \frac{-\frac{Q_C(0)}{C_{of}(1+\alpha)}}{1-\frac{Q_C(0)}{C_{of}(1+\alpha)}\left(f_BB+\frac{\mu}{2v_{sat}L_e}\right)} \times \frac{1}{\frac{1}{2}+\left[\frac{1}{4}+\frac{f_BB\frac{Q_C(0)}{C_{of}(1+\alpha)}}{\left\{1-\frac{Q_C(0)}{C_{of}(1+\alpha)}\left(f_BB+\frac{\mu}{2v_{sat}L_e}\right)\right\}^2}\right]^{\frac{1}{2}}} \tag{2.35}$$

Here L_e and $V_{DS(eff)}$ are unknowns; the saturation current is fully characterized by (2.35) and the channel-length modulation analysis for L_e, which is described next.

2.3.4.4 Channel-Length Modulation For Strong Inversion

In the analysis of the channel-length modulation [Vee88], we solve a differential equation which is subject to 2-D Gauss's law in the high

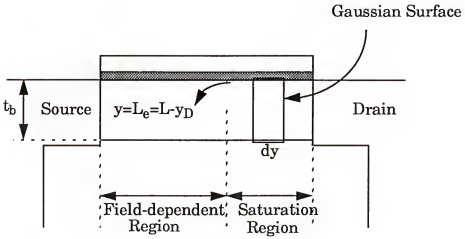


Fig. 2.6 Schematic cross section along the channel in saturation showing the field-dependent and saturation region with t_b and Gaussian surface for the channel-length modulation analysis.

longitudinal-field region where the carriers drift with their saturation velocity. The continuity of the current dictates the aerial charge density in this region is constant at $Q_C(L_e)$. To obtain a differential equation for $\Delta\Psi_{sf}(y)$, we apply Gauss's law for the V_{DS}-induced incremental field and charge in a subregion of length dy, as illustrated in Fig. 2.6:

$$\Delta Q_C(y)\,dy = \varepsilon_s \Delta E_x(t_b, y)\,dy - \varepsilon_{ox}\Delta E_{ox}(y)\,dy + \varepsilon_s dy \int_0^{t_b} \frac{d}{dy}\Delta E_y\,dx . \tag{2.36}$$

With (2.9), we can express the left-hand side of (2.26) as

$$\Delta Q_C(y) = \Delta Q_C(L_e) = C_{of}(1+\alpha)\Delta\Psi_{Sf}(L_e) - \frac{\varepsilon_s t_b \eta}{2} . \tag{2.37}$$

The DICE analysis allows the incremental change of the electric field at $x=t_b$ to be expressed in terms of $\Delta\Psi_{sf}$:

$$\Delta E_x(t_b, y) = \frac{\Delta\Psi_{sf}(y)}{t_b} + \frac{\eta t_b}{2} . \tag{2.38}$$

The second term on the right-hand side of (2.36) can be expressed as

$$\varepsilon_{ox}\Delta E_{ox}(y) = \varepsilon_{ox}\frac{\Delta\Psi_{ox}(y)}{t_{ox}} = -C_{of}\Delta\Psi_{sf}(y) . \tag{2.39}$$

For the integral term on the right-hand side of (2.36), we approximate

$$\varepsilon_s \int_0^{t_b} \frac{d}{dy}\Delta E_y\,dx = -\varepsilon_s \frac{d^2}{dy^2}\int_0^{t_b}\Delta\Psi\,dx \cong -\frac{\varepsilon_s t_b}{2}\frac{d^2}{dy^2}(\Delta\Psi_{sf}) \tag{2.40}$$

because $\Delta\Psi_{sb}(y) = 0$.

Substituting (2.37) to (2.40) into (2.36), we get

$$\frac{d^2}{dy^2}\Delta\Psi_{sf} = 2C_{of}(1+\alpha)\left(\frac{\Delta\Psi_{sf}(y)-\Delta\Psi_{sf}(L_e)}{C_b t_b^2}\right)+2\eta. \tag{2.41}$$

The boundary conditions for (2.41) are

$$\Delta\Psi_{sf}(L_e) = V_{DS(eff)},$$

$$\Delta\Psi_{sf}(L) = V_{DS},$$

and

$$\left.\frac{d}{dy}\Delta\Psi_{sf}\right|_{y=L_e} = \frac{2v_{sat}}{\mu_{eff}}. \tag{2.42}$$

With (2.42), the solution of (2.41) is

$$V_{DS}-V_{DS(eff)} = \frac{2v_{sat}l_c}{\mu_{eff}}\sinh\left(\frac{L-L_e}{l_c}\right)+2\eta l_c^2\left[\cosh\left(\frac{L-L_e}{l_c}\right)-1\right] \tag{2.43}$$

where $l_c = \sqrt{\frac{\varepsilon_s t_b}{2C_{of}(1+\alpha)}}$.

Now (2.43) with (2.35) implicitly models the channel-length modulation and saturation characteristics. For a very short channel length (i.e., deep-submicron channel length), the entire channel can become the saturation region [Fos93b]. In this case, the channel-length modulation is modeled such that $L_e \cong 0$.

2.3.5 Weak-Inversion Current

In weak inversion where $V_{GfS} < V_{TW}$, the channel current is predominantly diffusion [Sze81]:

$$I_{CH} = WV_{TH}\mu_n\frac{\partial}{\partial y}Q_c(y) \tag{2.44}$$

where μ_n is an average channel mobility defined as

$$\mu_n = \frac{\mu_{n0}}{1+\theta E_{Xf}}, \tag{2.45}$$

and

$$-Q_C(y) \cong \frac{qn_i^2 V_{TH}}{N_B E_{Xf}} \exp\left(\frac{\Psi_{sf} - V(y)}{V_{TH}}\right) \tag{2.46}$$

where V(y) is the quasi-Fermi level separation at y due to V_{DS}. Integrating (2.44) along the channel gives the expression for the channel current in terms of Q_C in (2.46). For short-channel devices, however, the surface potential is modulated along y because of the source and drain junctions in the two-dimensional structure. This modulation is difficult to distinguish from the charge-sharing effect; but it can be effectively modeled as channel-length modulation in weak inversion [Gre93b]. Thus we assume that (2.44) and (2.46), with constant Ψ_{sf}, are only applicable to the central portion of the channel ($y_S \le y \le L - y_D$) where Q_C can be well defined without 2-D effect from source and drain. Then with (2.46), integrating (2.44) from y=y_S to y=L-y_D yields

$$I_{CH} \cong \frac{W}{L_e}\frac{qn_i^2\mu_n}{N_B}\frac{V_{TH}^2}{E_{Xf}}\exp\left(\frac{\Psi_{sf}}{V_{TH}}\right)\left[1-\exp\left(-\frac{V(L-y_D)}{V_{TH}}\right)\right]. \tag{2.47}$$

where L_e=L-y_S-y_D.

The total surface potential, Ψ_{sf}, is composed of Ψ_{sf0} and $\Delta\Psi_s$ where $\Delta\Psi_s$ is given by (2.10). Ψ_{sf0} is calculated using (2.1) with $Q_C \cong 0$:

$$\Psi_{sf0} = \frac{1}{1+\alpha}\left[V_{GfS} - V_{FB}^{f} + \frac{Q_b}{2C_{of}} + \alpha V_{BS}\right]. \tag{2.48}$$

The average transversal field is defined as $\bar{E}_{xf} = \bar{E}_{xf0} + \Delta\bar{E}_{xf}$ where $\bar{E}_{xf0}$ is given by (2.16) and $\Delta\bar{E}_{xf}$, resulting from DIBL, is expressed as

$$\Delta E_{xf} = -\frac{C_{of}}{\varepsilon_s}\Delta\Psi_{sf} = -\frac{\eta t_b}{2(1+\alpha)}. \tag{2.49}$$

In (2.47), the only unknown thus far is $L_e = L\text{-}y_S\text{-}y_D$. To obtain the expression for L_e, we again solve the 2-dimensional Gauss's law near drain similar to the strong-inversion channel-length modulation analysis, but with different boundary conditions which are consistent with the diffusion assumption [Gre93b]. Starting from (2.36) and neglecting the channel charge, we obtain

$$\frac{d^2}{dy^2}\Delta\Psi_{sf} = \frac{1}{l_c^2}\Delta\Psi_{sf} + \eta, \tag{2.50}$$

the boundary conditions to which are

$$\Delta\Psi_{sf}(L) = V_{DS},$$

$$\Delta\Psi_{sf}(L - y_D) = \frac{\varepsilon_s t_b \eta}{2C_{of}(1+\alpha)},$$

and

$$\left.\frac{d}{dy}\Delta\Psi_{sf}\right|_{y = L - y_D} \cong 0. \tag{2.51}$$

Solving (2.50) with (2.51), we get

$$V_{DS} = 2\eta l_c^2 \cosh\left(\frac{y_D}{l_c}\right) - \eta l_c^2. \tag{2.52}$$

We can solve (2.43) for y_D, which is independent of V_{DS}. This independence implies $y_D = y_S$ for any V_{DS} because the symmetry of the MOSFET device underlies $y_D = y_S$ for $V_{DS} = 0$. Then the modulated channel length, L_e is defined as

$$L_e \equiv L - 2y_D \triangleq L - 2l_c \ln\left(\frac{L^2}{2l_c^2}\right). \tag{2.53}$$

2.3.6 Moderate Inversion - Weak and Strong Inversion Linking

The linking of strong- and weak-inversion current is done via cubic-spline interpolation across the moderate inversion region ($V_{TW}<V_{GfS}<V_{TS}$). The moderate-inversion boundaries defined previously are physical, and hence the continuity of the channel current and transconductance is ensured. To define the cubic spline, the current and the transconductance at each boundary are calculated from the respective formalism. The derivative in weak inversion is analytically evaluated while in strong inversion it is numerically determined. The spline is applied to the logarithm of the channel current in the moderate-inversion region:

$$\ln(I_{CH}) = \alpha_0 + \alpha_1 (V_{GfS} - V_{TW}) + \alpha_2 (V_{GfS} - V_{TW})^2 + \alpha_3 (V_{GfS} - V_{TW})^3 \tag{2.54}$$

where

$$\alpha_0 = \ln[I_{CH}(V_{TW})]$$

$$\alpha_1 = \frac{d\ln[I_{CH}(V_{TW})]}{dV_{GfS}}$$

$$\alpha_2 = \frac{3\ln[I_{CH}(V_{TS})/\ I_{CH}(V_{TW})]}{(V_{TS}-V_{TW})^2} - \frac{\frac{d\ln[I_{CH}(V_{TS})]}{dV_{GfS}} + 2\frac{d\ln[I_{CH}(V_{TW})]}{dV_{GfS}}}{V_{TS}-V_{TW}}$$

$$\alpha_3 = \frac{2\ln[I_{CH}(V_{TS})/\ I_{CH}(V_{TW})]}{(V_{TS}-V_{TW})^3} + \frac{\frac{d\ln[I_{CH}(V_{TS})]}{dV_{GfS}} + \frac{d\ln[I_{CH}(V_{TW})]}{dV_{GfS}}}{(V_{TS}-V_{TW})^2}. \qquad (2.55)$$

2.4 LDD and LDS

2.4.1 Ohmic Voltage Drop in LDS/LDD

The ohmic voltage drop across the LDS/LDD region can be modeled simply as a bias-independent resistance. Assuming the symmetry between the LDD and the LDS and the uniform current in the regions, the resistance in LDS/LDD is calculated as

$$R_{LDS/D} = \frac{L_{LDS/D}}{qWt_{LDS}\mu_{LDS/D}N_{LDS/D}} \qquad (2.56)$$

where $L_{LDS/D}$ is the length of LDS/LDD, and t_{LDS} is the thickness of the region. t_{LDS} is defined as t_f - t_{halo} where t_f is the total SOI film thickness and t_{halo} is the thickness of the halo region. $\mu_{LDS/D}$ in (2.56) is the majority carrier mobility dependent on $N_{LDS/D}$ [Mul77]. The defined resistance is lumped to

the source/drain parasitic resistance. Therefore, an iterative scheme is not needed to characterize this ohmic drop.

2.4.2 Non-Ohmic Voltage Drop in LDD

According to the maximum longitudinal electric field (at y=L) in the channel, a depletion region can possibly be induced in the LDD. Hence the voltage drop cannot be modeled as a simple resistor. Assuming uniform current, I_{DS}, in the LDD and no electric field modulation by the carriers [Cho91a], the non-ohmic voltage drop in the LDD is modeled considering three possible conditions, as shown in Fig. 2.7.

The constant field region in Fig. 2.7 represents a neutral part of the LDD which provides a simple ohmic drop, and the graded region reflects a depleted part characterized by a constant slope from Poisson's equation:

$$\frac{d|E|}{dy} = -\frac{qN_{LDD}}{\epsilon_S} = -S. \tag{2.57}$$

The analysis of the LDD is linked with the channel analysis by the maximum electric field,

$$E_m = \frac{V_{DS} - V_{LDD} - V_{DS(eff)}}{l_c} \tag{2.58}$$

as implied by the channel-length modulation analyses. In (2.58), $V_{DS(eff)}$ is defined by (2.35) in strong inversion, is zero in weak inversion, and is calculated via linear interpolation in moderate inversion using strong and weak inversion boundary values. When $V_{DS} - V_{LDD} < V_{DS(eff)}$, it is assumed that E_m is small enough to ensure that only 'case 1' in Fig. 2.7 is possible. Otherwise

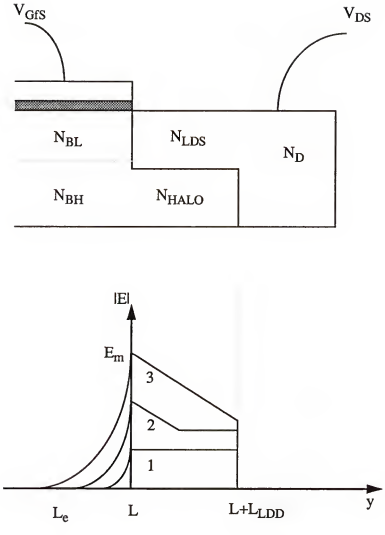


Fig. 2.7 The representation of the drain structure including the LDD, and three possible cases of the drain field distribution in the LDD.

all three cases are possible, depending on the resistance per unit length in the neutral LDD defined as

$$r_D = \frac{1}{qWt_{LDS}N_{LDD}\mu_{LDD}}. \tag{2.59}$$

If $E_m \le I_{DS}r_D$, then 'case 1' applies. If E_m exceeds $I_{DS}r_D$, then 'case 2' and 'case 3' can apply. The boundary between 'case 2' and 'case 3' is defined by the condition $E_m - I_{DS}r_D = SL_{LDD}$, which can be derived from Fig. 2.7. After the electric field condition is determined, the non-ohmic voltage drop across the LDD for each case is calculated:

$$\text{Case 1: } V_{LDD} = 0 \qquad \text{if } E_m < Ir_D$$

$$\text{Case 2: } V_{LDD} = \frac{(E_m - I_{DS}r_D)^2}{2S} \qquad \text{if } E_m > I_{DS}r_D, \text{ and } E_m - I_{DS}r_D < SL_{LDD}$$

$$\text{Case 3: } V_{LDD} = \frac{L_{LDD}(2E_m - SL_{LDD})}{2} - I_{DS}r_D L_{LDD} \qquad \text{if } E_m > I_{DS}r_D, \text{ and } E_m - I_{DS}r_D > SL_{LDD} \tag{2.60}$$

Note that V_{LDD} is calculated based on results of the channel analysis and the channel analysis is in turn dependent on V_{LDD} via diminished intrinsic drain bias ($V_{DS} \rightarrow V_{DS}$-$V_{LDD}$). Hence an iterative numerical method is adopted to physically link the two regional analyses.

2.5 Impact-Ionization Current

Impact-ionization current is characterized by evaluating the multiplication factor, (M-1). Only weak impact ionization ($M \approx 1$) is relevant

(for example, to model the BJT-induced breakdown [You88]); thus the characterization is done as a post-processing after the channel analysis. The multiplication factor is calculated by integrating the impact-ionization rate over the high-field region, including the high-field portion of the channel and the LDD:

$$(M-1) = \int_{(L-y_D)}^{L} \alpha dy + \int_{L}^{(L+L_{LDD})} \alpha dy. \tag{2.61}$$

For evaluating (M-1) in (2.61), we adopted a recently developed nonlocal model [Kri94], in which (M-1) is dependent on the carrier temperature rather than the longitudinal electric field directly. The results of the channel analysis, i.e., y_D and the electric field distribution in the high-field region, provide physical information needed for nonlocal carrier temperature characterization. Then the total generation current due to impact ionization is given as

$$I_{Gi} = (M-1)(I_{CH} + I_T) \tag{2.62}$$

where I_T is the parasitic-BJT current described next.

2.6 Parasitic-BJT Transport Current

The parasitic (lateral) BJT in the MOSFET structure is important to predict the loss of gate control [You88] and the transient-induced BJT action [Pel95]. For the injection condition in the neutral high-doped body, we assume low-injection. However, the injection condition at the source-body junction in the upper depletion region cannot be clearly defined due to the uncertain amount of majority carriers that depends on the degree of depletion; the

situation is similar to that in the FD device. Thus the BJT transport current is modeled as comprising components from both the depleted body and the neutral body:

$$I_T \cong W\,(J_{D0} + J_{N0})\left[\exp\left(\frac{V_{BS}}{V_{TH}}\right) - \exp\left(\frac{V_{BD}}{V_{TH}}\right)\right] \tag{2.63}$$

where J_{D0} and J_{N0} are structure-dependent constants, given as

$$J_{N0} = \frac{q n_i^2\,(t_f - t_b)\,D_n}{N_{BH} L_e} \tag{2.64}$$

and

$$J_{D0} = \frac{q n_i^2\,t_b \overline{D}_n}{\int_0^{(L-y_D)} p\left(V_{BS}, V_{BD}, V_J^{MIN}\right) dy}, \tag{2.65}$$

where V_J^{MIN} is a fitting parameter that defines the background majority-carrier density in the depleted body, and D_n and $\overline{D}_n$ are the diffusivity in the respective regions. When there is a halo region (i.e., p^+ region under LDS in Fig. 2.1), the minority-carrier injection into the neutral region tends to be suppressed due to the higher doping density, and thus the denominator in (2.64) is augmented by $N_{halo} L_{LDS}$.

2.7 Recombination/Thermal-Generation Current

The characteristics at the source/drain-body junctions in the NFD device are critical in predicting the floating-body effects. Due to the transient floating-body charging effect, we cannot afford the assumption that either a

forward- or a reverse-bias condition prevails at respective junctions, as was possible in FD model [Cho91a]. Therefore, it is modeled separately for recombination component and for thermal generation component, characterizing forward- and reverse-bias junction characteristics respectively. The weighted sum of the two components achieves the symmetric characteristics for both junctions and the smooth transition near $V_{BS/D} = 0$.

The recombination current in the FD device occurs predominantly in the quasi-neutral source region because of carrier separation due to the transverse field in the body [Fos93a] that suppresses the recombination in the space-charge region. This is not the case in the NFD device, however, due to the neutral region in the body. The simulated recombination current versus the quasi-Fermi level separation at the source junction ($\sim V_{BS}$), obtained from a 2-dimensional numerical device simulator (PISCES [PIS89]), is shown in Fig. 2.8. It shows a typical non-ideal junction characteristics in contrast to the FD counterpart. Thus the pertinent recombination current associated with the source-body junction is modeled generally as [Sze81]

$$I_R \cong WJ_{R0}\left[\exp\left(\frac{V_{BS/D}}{mV_{TH}}\right) - 1\right] + WJ_{S0}\left[\exp\left(\frac{V_{BS/D}}{V_{TH}}\right) - 1\right] \qquad (2.66)$$

where J_{R0} is a structure-dependent constant dependent on recombination life-time and $m\ (\geq 1)$ is a nonideality factor; the second component in (2.66) represents the quasi neutral-region recombination, and physically limits $V_{BS/D}$.

The thermal generation current is assumed to derive from the depletion region in the body [Fos93a]:

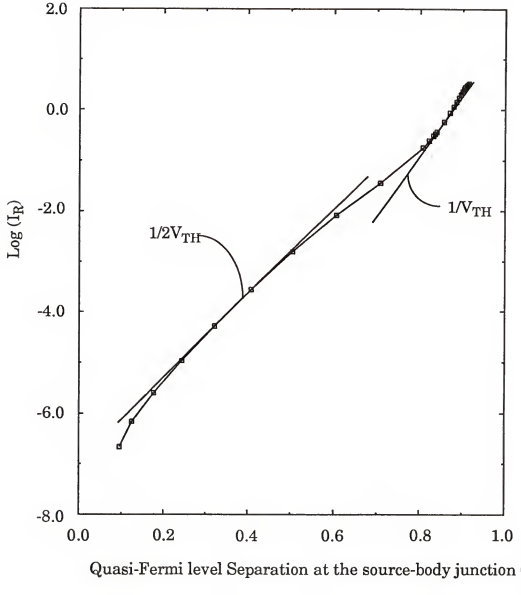


Fig. 2.8 The recombination current in the NFD device as a function of quasi-Fermi level separation at the source junction. The slope is approximately $1/2V_{TH}$ for low forward bias and $1/V_{TH}$ high forward bias.

$$I_{Gt} \cong \frac{qn_i W t_b L}{\tau_G}\left[\exp\left(\frac{V_{BS/D}}{V_{TH}}\right)-1\right] \tag{2.67}$$

where τ_G is the generation lifetime.

Now the weighted sum of (2.66) and (2.67) defines I_{RGt}, which characterizes both source and drain junction for all bias conditions:

$$I_{RGt} = \frac{\frac{J_{Gt0}}{J_{R0}}\exp\left(\frac{V_{BS/D}}{V_{TH}}\right)I_R + I_{Gt}}{1+\frac{J_{Gt0}}{J_{R0}}\exp\left(\frac{V_{BS/D}}{V_{TH}}\right)} \tag{2.68}$$

where $J_{Gt0} = qn_i t_b L/\tau_G$. The weighting factor, $(J_{Gt0}/J_{R0})\exp(V_{BS/D}/V_{TH})$, approximately weights I_R or I_{Gt} while effectively smoothing $I_{RGt}(V_{BS/D})$ when $V_{BS/D} \sim 0$.

2.8 Charge Modeling

The MOSFET charge components in Q_{Gf}, Q_D, and Q_S are modeled following [Vee89] in strong inversion. In weak inversion, the models are simplified by neglecting the channel charges. Then a linear interpolation of the intrinsic terminal charges defined at strong- and weak-inversion boundaries gives the descriptions for moderate-inversion charges:

$$Q_X^{mod} = Q_X^{wk} + \left(Q_X^{str} - Q_X^{wk}\right)\frac{V_{GfS} - V_{TW}}{V_{TS} - V_{TW}} \tag{2.69}$$

where X denotes Gf (front gate), S (source), and D (drain).

In a scaled NFD/SOI MOSFET, parasitic-BJT charges can be

important due to the possible triggering of parasitic BJT in transients [Pel95]. Thus Q_S and Q_D must include these charges. Q_{Gb} is defined from Gauss's law applied to the back-gate structure under the body region, and is augmented by the charge component associated with the substrate depletion under the source/drain due to a low-doped substrate whose mirror charges are also added to Q_S and Q_D.

Then the charge neutrality of the device defines Q_B, which also includes BJT charges:

$$Q_B = -(Q_S + Q_D + Q_{Gf} + Q_{Gb} + Q_{ff} + Q_{fb}) \tag{2.70}$$

where Q_{ff} and Q_{fb} are front and back fixed interface charges, respectively.

The terminal charging/discharging currents are characterized using the quasi-static approximation:

$$\frac{dQ_i}{dt} = \sum_j \frac{\partial Q_i}{\partial V_j}\frac{dV_j}{dt} \tag{2.71}$$

where i=S, D, Gf, Gb, B and j=DS, GfS, GbS, BS. Note that (2.71) defines non-reciprocal capacitances and transcapacitances that physically reflect the charge dynamics of the five-terminal device.

2.8.1 Front-Gate Charge

Gauss's law defines the front-gate charge,

$$Q_{GfS} = WC_{of}\int_0^L \left(V_{GfS} - \Phi_{ms}^f - \Psi_{sf} \right) dy . \tag{2.72}$$

In the triode region of strong inversion, (2.72) is integrated from y=0 to y=L to yield

$$Q_{Gf} = WLC_{of}\left[V_{GfS} - \Phi_{ms}^{f} - \Psi_{sfS} - \frac{V_{DS}}{2} + \frac{V_{DS}^{2}(1+s)(1+\alpha)}{-\frac{Q_C(0)}{C_{of}}\left[12 - \frac{6}{u}\right]}\right] \quad (2.73)$$

where Φ_{ms}^{f} is the front gate work-function difference. We also define $s=\mu_{eff}V_{DS}/2v_{sat}L$ and $u=-Q_C(0)/C_{of}(1+\alpha)V_{DS}$: $Q_C(0)$ is given by (2.1) and (2.9). When the device is in saturation, L is replaced by L_e (=L-y_D) and V_{DS} by $V_{DS(eff)}$ in (2.73), and then (2.73) is augmented by a term to reflect the charge in the saturation region from y=L-y_D to y=L:

$$Q_{Gf}^{sat} = WC_{of}(L-L_e)\left(V_{GfS} - \Phi_{ms}^{f} - \Psi_{sfS} - V_{DS(eff)}\right)$$

$$-WC_{of}(L-L_e)\frac{2l_c^{2}v_{sat}}{\mu_{eff}}\left(\cosh\left(\frac{L-L_e}{l_c}\right) - 1\right). \quad (2.74)$$

In weak inversion, the gate charge is calculated only at the center region of the channel whose length is L_e (= L - $2y_D$) as

$$Q_{Gf} = WL_eC_{of}\left(V_{GfS} - \Phi_{ms}^{f} - \Psi_{sf}\right). \quad (2.75)$$

The charge when the front surface is accumulated is also accounted for using (2.75) by pinning Ψ_{sf} to $V_{BS} + V_{TH}$ when it is less than $V_{BS} + V_{TH}$.

2.8.2 Drain and Source charges

In weak inversion, the channel charges are neglected. Thus the channel-charge component of drain and source charges are constant and defined as zero

$$Q_{D(CH)} = Q_{S(CH)} = 0. \tag{2.76}$$

In strong inversion, we partition the channel charge between source and drain following the scheme in [Vee88]. In triode region, the channel charge is

$$Q_{CH} = WC_{of}\int_0^L Q_C(y)\,dy$$

$$= -WLC_{of}(1+\alpha)V_{DS}\left[\frac{2z^3-(z-1)^3}{3}\frac{}{2z-1} + (u-z)\right] \tag{2.77}$$

where $z = u - (I_{DS}/2v_{sat}W)/C_{of}(1+\alpha)V_{DS}$. Drain and source charges are then computed as

$$Q_{D(CH)} = W\int_0^L \frac{y}{L} Q_C(y)\,dy$$

$$= -WLC_{of}(1+\alpha)V_{DS}\left[\frac{2}{3}\frac{(z-1)^3}{2z-1} + \frac{4}{15}\frac{z^5-(z-1)^5}{(2z-1)^2} + \frac{(u-z)}{2}\right], \tag{2.78}$$

$$Q_{S(CH)} = Q_{CH} - Q_{D(CH)}. \tag{2.79}$$

In saturation region, L is replaced by L_e and V_{DS} by $V_{DS(eff)}$ in (2.77), (2.78) and (2.79). The high-field region charges then must be added to these equations which are characterized as follows:

$$Q_{CH}^{sat} = W(L-L_e)Q_C(L_e) \tag{2.80}$$

$$Q_{D(CH)}^{sat} = W\frac{\left(L^2-L_e^2\right)}{2L}Q_C(L_e) \tag{2.81}$$

$$Q_{S(CH)}^{sat} = Q_{CH}^{sat} - Q_{D(CH)}^{sat}. \tag{2.82}$$

Source/drain BJT charges, $Q_{S(BJT)}$ and $Q_{D(BJT)}$, are calculated by characterizing the injected minority carriers across S/D-B junctions:

$$Q_{S/D\,(BJT)} = -\frac{qn_i^2 WL_{diff} t_f}{N_{DS\,(eff)}}\left[\exp\left(\frac{V_{BS/D}}{V_{TH}}\right)-1\right]$$

$$-\frac{qn_i^2 WL_e t_b}{2p_{S/D}\,(0)}\left[\exp\left(\frac{V_{BS/D}}{V_{TH}}\right)-1\right]$$

$$-\frac{qn_i^2 WL_e\,(t_f - t_b)}{2N_{BH}}\left[\exp\left(\frac{V_{BS/D}}{V_{TH}}\right)-1\right], \qquad (2.83)$$

where $N_{DS(eff)}$ is the effective doping in the source and drain which is smaller than the actual doping density due to high doping effects [Fos81], and L_{diff} is the effective diffusion length in the quasi-neutral source/drain. The first term in (2.83) represents the BJT charge in the quasi-neutral source and drain, whereas the next two terms represent those in the body region; $p_{S/D}(0)$ is the majority-carrier density at the junction in the low-doped body, which is consistent with I_T model, described previously.

The substrate-charge component in the drain/source charge is characterized by recognizing the parasitic MOS structure defined by high-doped source/drain, back-gate oxide, and substrate. Gauss's law is applied to the parasitic MOS capacitance to define this component:

$$Q_{S/D\,(DEP)} = A_{S/D} C_{ob}\left(-V_{GbS/D} - \Phi_{ms}^{SUB} - \Psi_{SUB}\right) \qquad (2.84)$$

where $A_{S/D}$ is the source/drain contact area, C_{ob} is the back-gate capacitance defined as ε_{ox}/t_{oxb}, and Φ_{ms}^{SUB} is the work-function difference in the MOS

structure defined by the doping densities in the source/drain and the substrate. Ψ_{SUB} is the surface potential in the substrate. It is calculated using depletion approximation to the MOS structure when $-V_{GbS/D} < V_{FB}^{SUB}$ where V_{FB}^{SUB} is defined as $\Phi_{ms}^{SUB} - Q_{fb}/A_{S/D}C_{ob}$. Otherwise, it is zero, reflecting the accumulated substrate interface.

The total source/drain charges are then defined as

$$Q_{S/D} = Q_{S/D(CH)} + Q_{S/D(BJT)} + Q_{S/D(SUB)} . \tag{2.85}$$

2.8.3 Back-Gate Charge

The Gauss's law implies the back-gate charge under the intrinsic body region as

$$Q_{Gb}^{int} = WLC_{ob}\left(V_{GbS} - \Phi_{ms}^{b} - V_{BS} \right) \tag{2.86}$$

where Φ_{ms}^{b} is back-gate work function difference. The total back-gate charge comprises (2.86) and the mirror charge of (2.84) for the substrate depletion:

$$Q_{Gb} = Q_{Gb}^{int} - (Q_{S(DEP)} + Q_{D(DEP)}) . \tag{2.87}$$

2.9 Conclusions

A physical model for the short channel NFD/SOI MOSFET has been developed. The model expands the TFA formalism [Vee88] to account for body-doping nonuniformities and LDD/LDS and halo regions, and to characterize the channel current in weak- and moderate-inversion. It also includes the generation current due to the nonlocal impact ionization and the thermal

generation, and the parasitic-BJT and recombination currents in the device which are important in the short-channel devices, especially when operating in the floating-body mode. Quasi-static charge modeling included for dynamic simulations completes the physical charge-based model, which is implemented in SOISPICE as described in the next chapter.

CHAPTER THREE
MODEL IMPLEMENTATION AND VERIFICATION

3.1 Introduction

This chapter gives a detailed description of the implementation of the NFD model described in the previous chapter into SOISPICE. Previous versions of SOISPICE [Fit91] were implemented only with a fully depleted (FD) SOI MOSFET model, and thus were insufficient to fully investigate contemporary technologies. With the addition of the NFD/SOI MOSFET model, we now have a very useful and complete tool for assessing the scaled SOI MOSFET technology, be it NFD or FD. The new NFD model is implemented analogously to the FD model; some of the parameters are used as flags for respective physical phenomena and structural hierarchy, and (trans)conductances and (trans)capacitances are evaluated numerically. However, the analogy does not go beyond the syntax as the newly implemented NFD model describes a completely different device structure having significantly different characteristics.

This chapter consists of two parts. First, Section 3.2 details the implementation of the new model. The network representation is presented, and the functions of the newly added/modified SPICE subroutines are defined. Then the algorithm that evaluates the intrinsic model elements is described.

The algorithm is a sequence of analyses that evaluate all currents and charges which correspond to a specified set of node voltages. The device-line and model parameters for the new model in SOISPICE are introduced and tabulated for their functions and meanings. In addition, numerical considerations and SPICE option parameters which can be used to obtain better results are mentioned whenever appropriate.

The second part, Section 3.3, describes tests of the newly implemented model, with predictions compared to two-dimensional device simulation results and measured data for wide range of channel lengths. These comparisons verify the model, but they involve careful considerations of parameter extraction. The new model is physics-based so that if reliable structural information is available, the extraction/optimization is mitigated as opposed to that with an empirically based model. The comparisons with the device simulations prove this point. Some model parameters in the comparisons with measurement data are evaluated by interpreting some of the measured characteristics using simple physical equations in the model. These comparisons show generally favorable results for two differently scaled technologies. The model predictions successfully show various floating-body effects as well as the effects of parasitic body resistance, which prove the predictability and the utility of the new tool in design studies of NFD/SOI MOSFET technology.

3.2 Model Implementation

3.2.1 Overall Structure

The network representation of the NFD/SOI MOSFET model is shown in Fig. 3.1. The device is modeled as a five-terminal one, with the option of the floating-body configuration. The floating-body option is chosen when the user does not specify the body node explicitly in the device line. Then, floating-body characteristics are simulated, accounting for all floating-body effects. Underlying this accounting is the non-zero V_{BS}, which is defined by the nodal equation $I_R=I_{Gi}+I_{Gt}$ in DC simulations but is influenced by dQ_B/dt in transient simulations. MODCHK, one of the SPICE subroutines, is revised so as to identify the body types and to reserve an internal node for floating-body configuration. The unlabeled internal node in the figure represents the enforcement of charge neutrality of the intrinsic charges.

To assemble the admittance matrix for the SPICE framework and evaluate the matrix elements properly for the NFD device shown in Fig. 3.1, four new subroutines (NFDFET, NFDID, NFDEC, and NFDMOD) were written, and 13 other SPICE subroutines were modified to accept the new device model. According to the type of analysis that SPICE requests, i.e., DC or transient, NFDFET initializes the variables used in the lower-level subroutines based on the previous values of the variables and on the dependence of the model components on them. The results from the lower-level subroutines using the initialized variables are then used to update and load the admittance matrix and the right-hand side current vector, as well as the state-

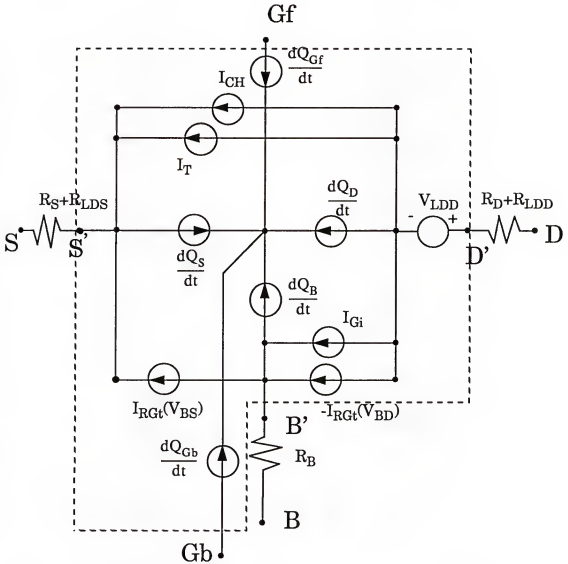


Fig. 3.1 The network representation of the new NFD/SOI MOSFET Model. Elements inside the broken line are intrinsic elements.

variable table for operating-point information. The elements in Fig. 3.1 are categorized into two groups (intrinsic and extrinsic), depending on the methods by which they are evaluated and loaded into the admittance matrix for SPICE iteration. The parasitic resistances and the ohmic resistances associated with the LDD/LDS, i.e., R_S, R_D, R_B, R_{LDS}, and R_{LDD}, are defined as extrinsic elements. The extrinsic elements are directly loaded into the matrix, once they are evaluated in MODCHK. When the LDD/LDS is part of the device, the LDS/LDD resistances are calculated in MODCHK from the model parameters that define the LDD/LDS structure, and lumped with the source/drain series resistances as represented in Fig. 3.1. The advantages of this method are twofold compared to SOISPICE-2 [Cho91a] in which the ohmic voltage drops across the LDS/LDD are accounted for by iterations inside the model subroutine: 1) we eliminate the iterative redundancy and save computation time, and 2) the effect of the intrinsic charge dynamics are correctly accounted for as the charging currents are flowing through the LDS/LDD resistances. NFDFET also adds the minimum shunting conductance in SPICE, GMIN, to the I_{RGt} branches in Fig. 3.1 which characterize the junction currents. This alleviates the convergence problem, commonly encountered in the reverse biased junction. The non-ohmic LDD voltage drop, V_{LDD}, is included as the intrinsic element because the analysis underlying it is iterative in the model routine.

NFDFET calls NFDID in DC simulations and NFDEC as well in transient simulations for evaluating intrinsic elements in the admittance

matrix. Because of the nonlinearity of the intrinsic elements, these elements need to be linearized for SPICE iterations. Thus when called by NFDFET, NFDID calls NFDMOD once for operating-point currents and voltages, and subsequently multiple times with slight perturbation of one of the four terminal voltage differences (relative to the source) each time to evaluate transconductances and transcapacitances numerically. Indeed, the implicit nature of our model hinders us from analytic evaluations of these derivatives, and hence this numerical evaluation scheme is required. The evaluated intrinsic variables are stored in the state-variable table as operating-point information. Using the (trans)capacitances from NFDID, NFDEC produces equivalent-circuit elements associated with the capacitances, including the overlap capacitances in transient simulations.

With a specific set of node voltages, NFDMOD will evaluate intrinsic currents and charges according to the NFD MOSFET model as described in Chapter 2 by following the algorithm which is detailed in the following subsection.

3.2.2 Algorithm of NFDMOD

The subroutine NFDMOD evaluates the intrinsic current elements, and charges as well, from a set of intrinsic nodal voltages, $V_{GfS'}$, $V_{D'S'}$, $V_{B'S'}$, and $V_{GbS'}$. The algorithm is flowcharted in Fig. 3.2, showing the iteration loop for the non-ohmic V_{LDD}. The first element to be calculated is I_{RGt} at both S/D-B junctions using $V_{B'D'}$ and $V_{B'S'}$. Strictly speaking, I_{Gt} (reverse-bias characteristics of I_{RGt}) is dependent on $V_{B'D'}$ minus non-ohmic V_{LDD}, but, the error

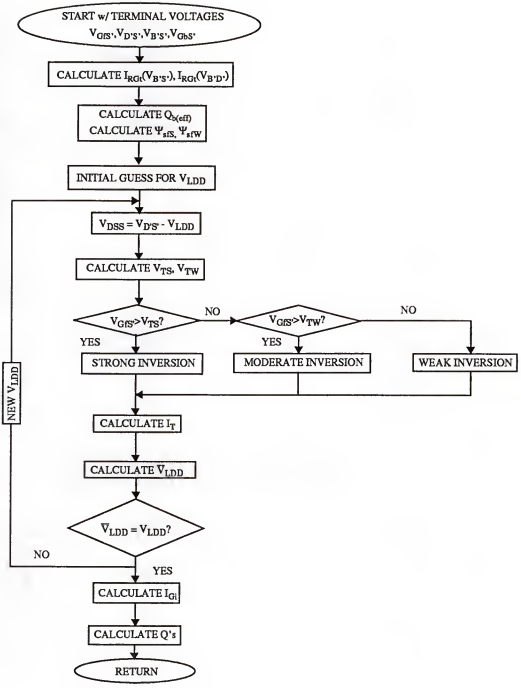


Fig. 3.2 The overall algorithm of NFDMOD.

from using $V_{B'D'}$ is negligible because a significant V_{LDD} can develop only when $V_{D'S'}$ is large which implies a significant reverse bias at this junction. Next, charge-sharing is characterized, yielding $Q_{b(eff)}$. Then we calculate Ψ_{sfS} and Ψ_{sfW} for the moderate-inversion boundaries. An initial guess for V_{LDD} is taken next, and the intrinsic $V_{D'S'}$ is diminished by V_{LDD} inside the V_{LDD} loop. V_{TS} and V_{TW} are then calculated. Because V_{TS} and V_{TW} must reflect the drain-induced effects (DIBL and DICE), which are functions of the intrinsic drain bias, they have to be re-evaluated for each iteration of V_{LDD}. Once the boundaries are determined, I_{CH} is evaluated according to the inversion condition of the channel depending on $V_{GfS'}$. I_T is evaluated next, and $I = I_{CH}+I_T$ is used to calculate $\overline{V}_{LDD}$, the non-ohmic drop in LDD, based on the maximum field derived from the channel analysis. Convergence is reached when V_{LDD} and $\overline{V}_{LDD}$ differ by less than the preset tolerance; otherwise a new guess is taken for next iteration. The iteration control is done by the Regula Falsi method [Mor83], which is a modified secant method, for efficient convergence. I_{Gi} is evaluated as post-processing. The intrinsic charges are evaluated last using converged current solutions and voltages.

3.2.3 NFD Device/Model Parameters and SOISPICE Input Description

The NFD model uses new device/model parameters for specifying a device to be simulated. The new device type is identified by a device name with Z prefix in SOISPICE. The external nodes follow the device name in the order of drain, front gate, source, back gate, and body. If a body node is not specified in the device line, SOISPICE internally generates a floating-body

Table 3.1

DEVICE LINE PARAMETERS

	Description	Units
L	Device Length (drawn)	m
W	Device Width (drawn)	m
BJT	BJT Flag	.
AS	Source Contact Area	m^2
AD	Drain Contact Area	m^2
NRS	Number of Squares of Source Contact	.
NRD	Number of Squares of Drain Contact	.
NRB	Number of Squares of Body Contact	.
IC	Initial Voltage Differences $V_{D'S'}$, $V_{GfS'}$, $V_{GbS'}$, $V_{B'S'}$	V

node which is not accessible externally. There are eight device-line parameters as listed in Table 3.1. They are device length (L), device width (W), BJT flag (BJT) which activates the parasitic BJT analysis, source and drain contact areas (AS, AD), and number of squares of source/drain and body contact (NRS, NRD, NRB). The BJT flag must follow the external nodes in the line if it is to be set. NRS, NRD, and NRB are used to evaluate the parasitic resistances associated with source, drain, and body respectively. When they are to be used, the sheet resistance model parameters RHOSD and RHOB must be specified. AS and AD are used to calculate the charges associated with substrate depletion under the source and drain. The last items of the device line are initial conditions which are used as first guesses for internal node voltages of the device during SPICE iterations. They are ordered as $V_{D'S'}$, $V_{GfS'}$, $V_{GbS'}$, and $V_{B'S'}$. Because V_{BS} defines the junction characteristics, a good initial guess for V_{BS} is highly recommended for convergence efficiency, especially in floating-body simulations.

The new NFD model is referred to by NNFDSOI or PNFDSOI in the .MODEL card, depending on the type of the model. There are 46 model parameters listed in Table 3.2 with their definitions and default values. A self-consistent relation is forced among work functions, flat-band voltages, and fixed interface charges; the unspecified of these three parameters is calculated physically in MODCHK. Some model parameters function as flags for the related effects and structures. The accountings for charge-sharing (ZETA), DICE/DIBL (ETA), channel-length modulation (LMOD), impact ionization

Table 3.2

MODEL PARAMETER LISTS

Name	Description	Units	Default
VFBF	Front-gate flatband voltage	V	0
VFBB	Back-gate flatband voltage	V	0
WKF	Front-gate work function difference	V	0
WKB	Back-gate work function difference	V	0
NQFF	Front oxide fixed-charge density	$\#/cm^2$	0
NQFB	Back oxide fixed-charge density	$\#/cm^2$	0
TOXF	Front-gate oxide thickness	m	1E-8
TOXB	Buried oxide thickness	m	5E-7
NSUB	Substrate doping density	$\#/cm^3$	1E15
NGATE	Poly Gate doping density	$\#/cm^3$	1E19
TF	Total silicon film thickness	m	2E-7
TB	Effective film thickness	m	1E-7
THALO	Halo thickness	m	0
NBL	Low body doing density	$\#/cm^3$	5E16
NBH	High body doping density	$\#/cm^3$	5E17
NHALO	Halo Doping density	$\#/cm^3$	0
UO	Zero-field mobility	$cm^2/Vsec$	700
THETA	Mobility degradation factor	m/V	1E-6
BFACT	Drain-bias averaging factor	V^{-1}	0.3
VSAT	Carrier saturation velocity	cm/sec	1E7

TABLE 3.2 --(Continued)

Name	Description	Units	Default
TPG	Type of gate material 0: metal (aluminum) -1: same as body type 1: opposite to body type	.	1
TPS	Type of substrate 0: metal (aluminum) -1: same as body type 1: opposite to body type	.	-1
NLDS	Doping density in LDD/LDS region	$\#/cm^3$	3E18
NDS	Doping density in the source/drain	$\#/cm^3$	5E19
LLDS	LDS region length	m	0
LLDD	LDD region length	m	0
ALPHA	Impact-ionization coefficient	.	0
BETA	Impact-ionization coefficient	m/V	0
CGFDO	Gate-drain overlap capacitance	F/m	0
CGFSO	Gate-source overlap capacitance	F/m	0
CGFBO	Gate-body overlap capacitance	F/m	0
RD	Extrinsic drain resistance	Ω/m	0
RS	Extrinsic source resistance	Ω/m	0
RB	Extrinsic body resistance	Ω	0
RHOSD	Source/drain-contact sheet resistance	Ω/❑	0
RHOB	Body-contact sheet resistance	Ω/❑	0
JRO	Junction recombination current coefficient	A/m	1E-10
M	Junction non-ideality factor	.	2
LDIFF	Effective diffusion length in S/D	m	0.1μm
VJMIN	Minimum pn junction voltage	V	0.7

TABLE 3.2 --(Continued)

Name	Description	Units	Default
TAUO	Carrier lifetime in lightly doped region	sec	1E-6
DL	Channel-length reduction	m	0
DW	Channel-width reduction	m	0
ZETA	Charge-sharing flag	.	1
ETA	DICE/DIBL flag	.	1
LMOD	Channel-length modulation flag	.	1

(ALPHA and BETA), LDS (LLDS), LDD (LLDD), and halo (THALO) can be turned off by setting the noted parameters to zero. LMOD must be set to 1 if impact-ionization is to be included in the simulations. Otherwise, the channel analysis will not provide the needed high-field region information. In order to represent a physically feasible device, TB and THALO must not be greater than TF. RD, RS, and RB override RHOSD and RHOB when they are specified simultaneously. A group of frequently encountered quantities in characterization are calculated using the model parameters as the intermediate parameters in MODCHK, and stored in separate locations for later reference.

An example input circuit file including an NFD device is shown in Fig. 3.3, listing all parameters. Optional device-line parameters are parenthesized. The state variables such as internal voltages, terminal charges, and physical device characteristics can be accessed as output variables using the .PRINT and the .PLOT cards in SPICE with the format, '&(device_name, variable_name)', for DC as well as transient simulations.

Note that some elements in the model frequently deal with very small numbers during iterations. The convergence criteria in SPICE must be set properly to yield meaningful values. SPICE, in general, uses various .OPTION parameters to reset the acceptable error tolerance for convergence. The default values are not always adequate for small-geometry MOSFETs. For example, ABSTOL defaulted to 10^{-9} is too large for I_{Gt} or I_{CH} in very weak inversion, and CHGTOL defaulted to 10^{-14} is too large for the intrinsic charges because of the small geometry of the scaled devices. Thus it is sug-

```
* This program tests the soispice3 for NFD SOIMOSFET
.WIDTH IN=80 OUT=80
*VOLTAGE SOURCE
VCC1 5 0 2.0
VGB1 3 0 0
VGF1 1 0 2.0
VDUM1 5 21 DC 0

*Device line
ZN1 21 1 0 3 (0) (BJT) NTYPE L=0.2E-6 W=10E-6 (AD=0.0) (AS=0.0) (NRD=1) (NRS=1)
+(NRB=2) IC=0.1,-0.1,0,0.6

.OP
.DC VGF1 -1 2 0.02
.PRINT DC I(VDUM1)
.PRINT DC &(ZN1,QD) &(ZN1,LE) &(ZN1,VTS) &(ZN1,VBS)
.OPTION ACCT LIST NODE LIMPTS=2000 TNOM=25 VNTOL=1E-6 PIVTOL=1E-30
+ABSTOL=1E-15 GMIN=1E-15 CHGTOL=1E-16

*Model card
.MODEL NTYPE NNFDSOI
 + VFBF=0.0D0 VFBB=0.0D0 TOXF=0.005U TOXB=0.4U WKF=0.0D0
 + WKB=0.0D0 NQFF=2e10 NQFB=1e11 NSUB=1.0E14 NGATE=2.0E19
 + TF=.075U TB=0.033U THALO=.04U NBL=4E17 NBH=1.0E18
 + NHALO=1.0E19 UO=400.0 THETA=1.1E-6 VSAT=5.0E6 TPS=-1.0
 + TPG=1.0 NLDS=1E18 NDS=5E19 LLDS=0.0U LLDD=0.0U
 + ZETA=1.0 ETA=1.0 LMOD=1.0 ALPHA=1.5E6 BETA=7.0E6
 + BFACT=0.5 CGFDO=0 CGFSO=0 CGFBO=0.0 RD=30
 + RS=30 RB=1000 TAUO=1E-9 DL=0.0 DW=0.0 JRO=1E-10 M=1.6
.END
```

Fig. 3.3 An example input description for NFD device/circuit in SOISPICE

gested that ABSTOL of 10^{-15} and CHGTOL of 10^{-16} or less be specified in contemporary simulations. We also used GMIN for each diode branch (S-B and D-B) to alleviate the convergence problem. GMIN shunts the current of the branch it is attached to, and hence influences the accuracy of the simulation. Consequently, smaller GMIN is more desirable, but the effect of GMIN is minimum when GMIN is smaller than ABSTOL because SPICE simply ignores the any current less than ABSTOL. Thus, GMIN equal to ABSTOL should be specified in the .OPTION card.

3.3 Verification of the NFD/SOI MOSFET Model

3.3.1 Comparisons with Numerical Simulation

To get preliminary support for the model and to confirm its predictive capabilities, we compared DC results of SOISPICE device-level simulations with I-V characteristics obtained from numerical device simulations using MEDICI [Med92]. In contrast to empirical SPICE models, our NFD/SOI MOSFET model is physical and its parameters can be effectively defined from the device structure and physical models used by MEDICI; no rigorous parameter extraction/optimization is needed. Indeed, the MEDICI input file for mesh generation and doping specification is sufficient to generate the structural parameters for the NFD model. The doping densities of distinctive regions in Fig. 2.3 were all specified using the 'uniform' parameter in MEDICI, and thus each region is separated by a well defined doping density. TB for the SOISPICE model is determined as the high-low junction depth in

the MEDICI device description because of the assumed large ratio (10) between the doping levels in the intrinsic body. UO in the SOISPICE simulations is set to 700 $cm^2/Vsec$, corresponding to the tabulated concentration-dependent mobility model used in the MEDICI simulations for the body doping of $1x10^{17}/cm^3$. Also, the temperature-dependent saturation velocity model in MEDICI yields $1x10^7$ cm/sec using default parameters at 300K, and thus, VSAT in the SOISPICE input file is set accordingly. The lone discrepancy between the MEDICI model and our SOISPICE model is the way that the mobility degradation is characterized. Thus, THETA had to be tuned for best fit, and maintained as the channel length and other parameters are varied. Table 3.3 lists the standard model parameters used in the comparisons, which are defined corresponding to the MEDICI device description. The extrinsic resistances were set to zero in the SOISPICE simulations. JRO and M are not important in these comparisons because impact ionization is turned off, and the simulated device structure has a body contact. Also, the parasitic capacitances are irrelevant in DC simulations.

A wide variety of device structures, with n-channel lengths ranging from 1 μm down to 0.2 μm, were simulated successfully. First, the model was tested for varying L with the standard model parameters in Table 3.3. It is exemplified in Fig. 3.4 where the SOISPICE- and MEDICI-predicted current-voltage characteristics are plotted for both suprathreshold and subthreshold characteristics for L=0.5μm. The predicted I_D-V_{DS} characteristics for an L = 0.5 μm device in Fig. 3.4 are in good agreement.

Table 3.3

STANDARD MODEL PARAMETERS FOR COMPARISONS WITH MEDICI RESULTS

Name	Value
TOXF	20 nm
TOXB	400 nm
TF	0.3 μm
TB	0.1 μm
NBL	$1.0x10^{17}$ /cm^3
NBH	$1.0x10^{18}$ /cm^3
LLDD/LLDS	0.25 μm
NLDS	$1.0x10^{18}$ /cm^3
NDS	$5.0x10^{19}$ /cm^3
UO	700 cm^2/Vsec
VSAT	$1.0x10^7$ cm/sec
THETA	$1.1x10^{-6}$ cm/V
NGATE	$1.0x10^{19}$ /cm^3
TPG	1
TPS	-1

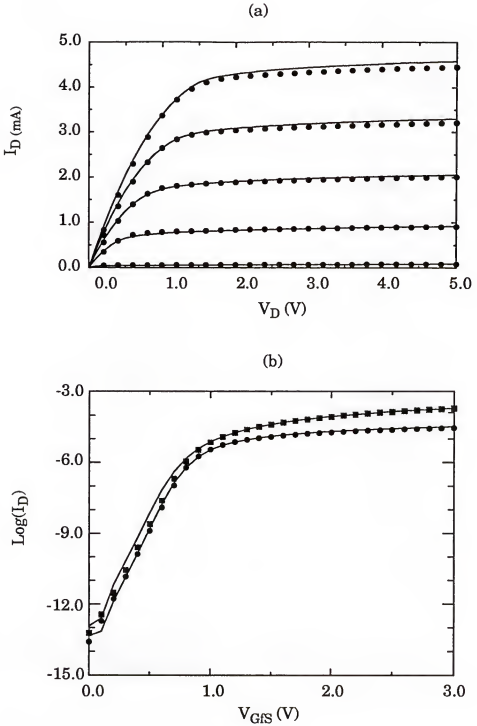


Fig. 3.4 MEDICI- (points) and model-(curves) predicted characteristics for an L = 0.5μm n-channel NFD/SOI MOSFET; (a) I_D-V_{DS} characteristics (V_{GfS} = 1 - 5 V); (b) Subthreshold characteristics (V_{DS} = 0.1, 3V). V_{BS} = 0V.

To further investigate the predictability of our model, the device structure was altered for a fixed L of 0.8 μm by varying model parameters: the device simulations were performed with 5, 10, and 20 nm of the gate oxide thickness and with 0.1μm and 0.05μm of high-low junction depth. SOISPICE reflected these structural variations excellently by resetting the corresponding model parameters (TOXF, TB) without additional optimization, as evidenced in Fig. 3.5. In the SOISPICE simulations of Fig. 3.5.(b), the value of TB is different from the high-low junction depth only when it is 0.05μm. This is because the depletion region under the front gate extends partly into highly doped region due to the low-doped region being too thin. In such cases, TB is tuned from the subthreshold slope which is related to the body depletion capacitances ($C_D = \varepsilon_s/TB$) as evident in Chapter 2. The calculated TB from this method is 0.056 μm when the high-low junction depth is 0.05 μm. This value is almost identical to the value obtained by applying the depletion approximation to the front MOS structure (0.058 μm) with the simulated doping profile. This idea is also used to determine TB for the measured I-V data. It is again stressed here that no additional optimization of the model parameters is needed once these parameters were tuned. These comparisons prove that our physical model is predictable and scalable, and has its utility in technology-CAD applications.

3.3.2 Comparisons with Measured Data

For additional verification, we compared model predictions with measured data taken from test devices from two different NFD/SOI technolo-

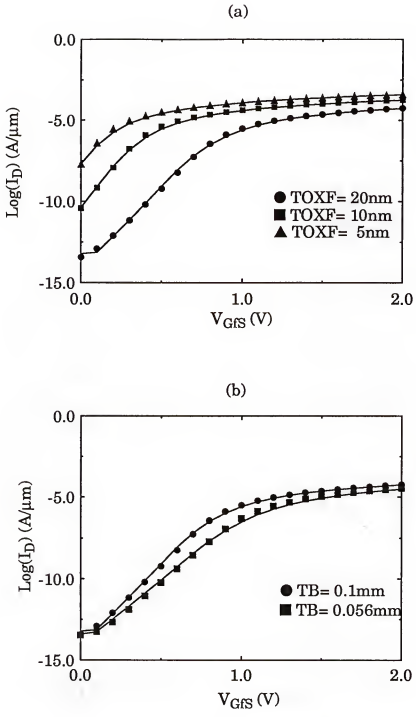


Fig. 3.5 MEDICI- (points) and model-(curves) predicted subthreshold characteristics for an L=0.8μm n-channel NFD/SOI MOSFET for varying (a) the gate oxide thickness (t_{oxf}) and (b) the effective film thickness (t_b). V_{DS} = 0.1V, V_{BS} = 0V.

gies, each having a wide range of channel lengths. The model parameters were determined in large part using device structural information: the front-gate thickness was available from the technology, and the LDD length was given by the spacer width. But some minor tuning was necessary because the available information was incomplete. For example, TB was calculated to be in accord with the measured subthreshold slope at low V_{DS} using (2.47) and (2.48). It was already noted in the previous subsection that the values thereby derived are nearly equal to the calculated maximum depletion widths for the doping profiles provided. In the real device, the doping profile in the body region will never be uniform nor known precisely, and hence it is reasonable and yet accurate to determine TB from the low-V_{DS} subthreshold slope. ALPHA and BETA for impact ionization were set constant at 2.45×10^6 and 1.92×10^6 respectively for the non-local model [Kri94]. The parameters responsible for the source-body junction (JRO and M) are determined from the high-V_{DS} subthreshold slope in the floating-body characteristics and the amount of shift in subthreshold curves with the drain bias. The constant ALPHA and BETA, with pretuned JRO and M, were consistent in predicting the kink effect on both subthreshold and suprathreshold characteristics. The carrier lifetime in the body, TAUO, is tuned by the saturation current before kinks.

The set of parameters thereby defined, reliably predicted I-V characteristics for channel lengths from 0.2μm to 1μm in two differently scaled technologies. In the first technology, the devices were fabricated on SIMOX wafers having a 400-nm buried oxide and a 300-nm silicon film, with LDS/

LDD and halo implants. The gate oxide is 20 nm thick, and the gate is n-type polysilicon. The devices have external body contacts, so both floating-body and body-tied (to source) characteristics could be measured. Results of the model application to this technology are exemplified in Fig. 3.6 where measured and predicted I_D-V_{DS} characteristics of an n-channel device with an effective channel length of 0.6 μm are plotted. The SOISPICE predictions for both body-tied and floating-body cases, done with a single set of model parameters, are generally good. The discernible floating-body effects, i.e., the premature breakdown, the saturation-current kink, and a low-V_{DS} current enhancement, are also modeled well. The latter effect has not been acknowledged well in previous works, but it is indeed significant. For example, at $V_{DS} = V_{GfS} = 1.5V$, the enhancement is near 50%. Note that there is evidence of premature breakdown even in the body-tied configuration; this is due to the finite body resistance in the body-source tie, which allows a nonzero V_{BS} when impact-ionization current is injected into the body (discussed in Chapter 4).

In the second technology, the devices were more aggressively scaled. They were fabricated on SOI substrates having 360 nm buried oxide and 75 nm silicon film, without LDS/LDD or halo. The gate oxide is 5 nm, and there is no body tie. The model application to this technology is exemplified by the measured and simulated current-voltage characteristics in Fig. 3.7 for a floating-body n-channel device with an effective channel length of 0.2 μm. The I_D-V_{DS} characteristics in Fig. 3.7(a) show good agreement, although some discrepancy is obvious at high I_D and V_{DS}, as it is in Fig. 3.6 to a less extent.

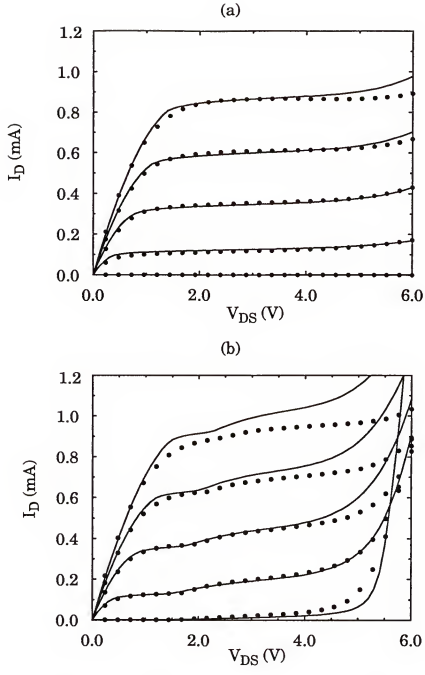


Fig. 3.6 Measured (points) and model-predicted (curves) I_D-V_{DS} characteristics of a W/L=2.4μm/0.6μm n-channel NFD/SOI MOSFET with LDD/LDS and halo: (a) tied-body characteristics, (b) floating-body characteristics. V_{GfS} = 1V - 5V.

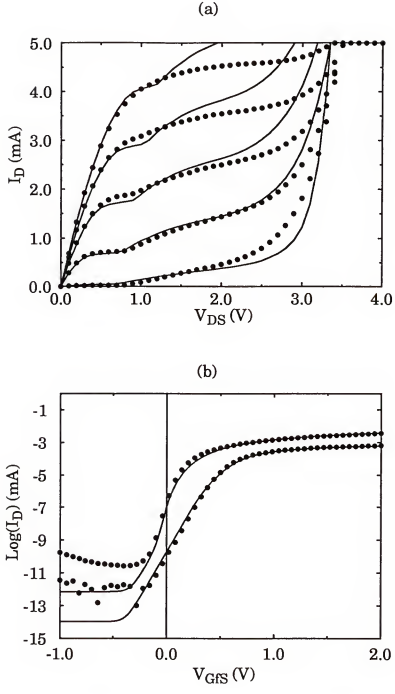


Fig. 3.7 Measured (points) and model-predicted (curves) (a) I_D-V_{DS} characteristics and (b) subthreshold characteristics of a W/L=10μm/0.2μm floating-body n-channel NFD/SOI MOSFET (without LDD/LDS). In (a) V_{GfS} = 0.5 - 2.5; in (b), V_{DS} = 0.1V, 2V.

These discrepancies can be attributed to the self-heating in the SOI MOSFET under DC bias producing substantive power dissipation [Su94], which is not accounted for in the model. The same floating-body effects noted in Fig. 3.6(b) are obvious in Fig. 3.7(a). The subthreshold I_D-V_{GfS} characteristics in Fig. 3.7(b), which also show very good agreement, however reveal another floating-body effect at high V_{DS}. The increased slope reflects a subthreshold kink, and this effect is also well predicted by our model.

3.4 Conclusions

The NFD/SOI MOSFET model, presented in Chapter 2, was implemented in SOISPICE. SOISPICE now contains both fully depleted and non-fully depleted SOI MOSFET models. The use of the model was explained: device/model parameters were listed and an example input file was presented. Because of the scaled sizes of the devices, it is suggested that pertinent option parameters in SPICE be specified with smaller values than default values. The new model was then tested against device simulation results and measured data. The model parameters could be evaluated from physical structures and models, either defined for the device simulation or available from technology information for measured device. Because this information was incomplete, tuning of some parameters was needed, based on the physical assumptions rendered by the model. SOISPICE predictions using our model showed favorable agreement with the data from the various devices, including numerous floating-body effects.

CHAPTER FOUR
THE EFFECT OF BODY RESISTANCE ON THE BREAKDOWN CHARACTERISTICS OF NFD/SOI MOSFETS

4.1 Introduction

One of the factors that limit the operating voltage in a CMOS circuit is the breakdown voltage (BV) of the constituent devices. Breakdown voltage of a MOSFET is a drain-source voltage above which the device loses gate control, and thus fails to operate as a switch. Usually breakdown voltage is required to be sufficiently larger than the operating voltage of circuit for reliable technology because of reliability and testability. Even though the operating voltage of bulk-silicon CMOS devices tends to decrease as the technology is scaled, it has been a severe limitation.

In short-channel SOI MOSFETs, the breakdown is premature and BV is lower than that of the bulk counterpart. This is due to the parasitic bipolar junction transistor (BJT), triggered by the impact-ionization generation current [You88], [Cho91b] when the body is floating. For a given generation current level, a positive feedback supported by finite body-source voltage eventually leads to the premature breakdown without a body contact to shunt off generated carriers. Intrinsic body-tied-to-source (BTS) structures, which drain the generated current out of the body, have been used in NFD/SOI MOSFETs to ameliorate this effect, but they do not completely eliminate

it [Hwa91] since the BTS strap (e.g., p^+ plug in n^+ source in Fig. 2.1) width must be limited relative to the effective device width. Due to the spacing between BTS straps and the finite conductivity of the thin body region, a non-zero resistance inherently exists between the body and the source regions. This body resistance, represented by a lumped value R_B which parallels the body-source pn junction, enables a BJT-induced premature breakdown, the mechanism of which is similar to that in the bulk MOSFET [Hsu82]; the generation current develops an ohmic drop across R_B, which acts as a forward bias at the source-body junction to turn on the parasitic BJT.

The breakdown conditions suggested in [Hsu82] can be misleading, however, because the holding voltage and the snapback voltage are indistinguishably referred to as the breakdown voltage, and the stated condition for breakdown, i.e., $(M-1)\beta = 1$, is independent of R_B. Hence, their model fails to address the effect of R_B which is apparent in the device characteristics such as in Fig. 3.6. To clarify this issue and to give physical insight regarding the efficacy and design of the BTS for NFD/SOI MOSFETs, we theoretically analyze in this chapter the breakdown characteristics and their dependence on R_B and other physical device parameters. The dependence of breakdown voltages on R_B is demonstrated in Section 4.2 using a 2-D device simulator [PIS89]. The effect of a halo is studied as well in this context. In Section 4.3, we derive simple but physical descriptions of the holding and snapback voltages. The derivations are aided and supported by SOISPICE simulations, using our newly implemented physics-based model for the NFD/SOI MOSFET. The

analyses help determine the efficacy of body ties in NFD/SOI MOSFETs in Section 4.4.

4.2 Breakdown Characteristics of an NFD Device with Halo

The halo implant is done after the LDS and LDD implant with a dopant of opposite type, forming a p-n junction around the LDS and LDD as indicated in Fig. 2.1. The doping level of the halo region is much higher than that of the body. The halo region affects many device characteristics: for example, it increases the punchthrough voltage and mitigates the short-channel threshold-voltage fall-off [Cod85]. Further, an increase of the BJT-induced breakdown voltage, BV, is experimentally evidenced in halo devices due to the high doping level. For a typical technology, the measured improvement of BV is about 0.7 volt in a halo device compared to the devices without it [Pol91]. But the mechanism which brings this improvement has not yet been studied adequately to facilitate an optimal design.

To gain some insights about the effects of the halo region, we use the 2-D device simulator, PISCES [PIS89]. Simulated body-tied device current-voltage characteristics are shown in Fig. 4.1. The the increase of BV in the halo device can be seen in the figure to be about 0.6 volts. These results coincide well with actual measured data. Figure 4.2(a) shows the structure of the device used for the simulations. Due to the two-dimensional constraints of the simulator, it can not accurately account for the three-dimensional flow of holes produced by impact ionization near the drain region in the double body-tied-

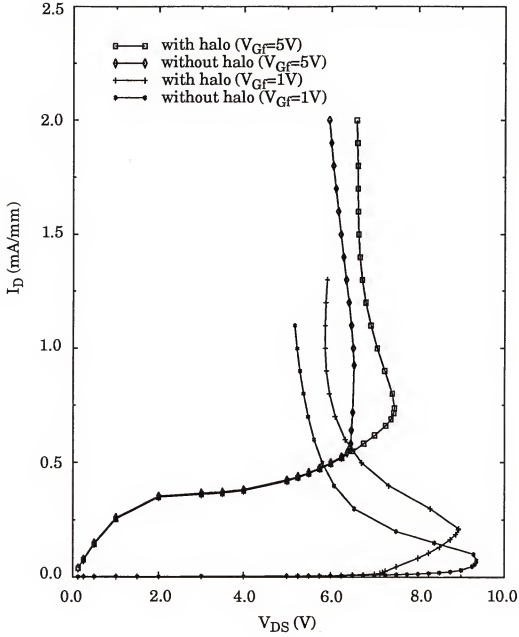
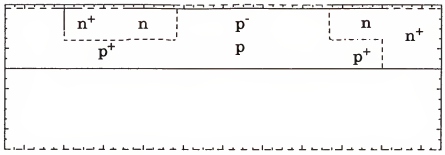


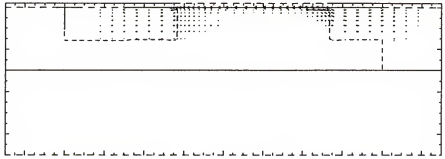
Fig. 4.1 PISCES-simulated I_D-V_{DS} characteristics of a typical n-channel BTS NFD/SOI MOSFET with and without halo implant. W/L=1μm/0.8μm

to-source (DBTS) structure [Hwa91]. We think however that the simulated structure is a good structural approximation of the actual halo-implanted DBTS structure, as implied by the simulation results. In the simulated structure the doping density in the body is 10^{17} cm^{-3} in the low-doped region and $2x10^{17}$ cm^{-3} in the high-doped region. For the structure with halo, the p-type doping is 10^{18} cm^{-3} under the source; the halo does not wrap around the source region in the simulated structure. For the structure without halo, the doping density under the source is identical with the doping density in the high-doped body. When the halo does not exist, the thickness of the LDD is same as the thickness of the Si film.

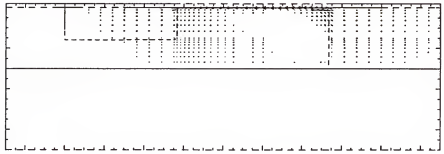
One effect of the halo is apparent in Fig. 4.2, where the predicted electron currents, represented by arrows in the body, are shown when the devices are biased at V_{DS} ~ BV (see Fig. 4.1). Considering that the arrows represent the relative magnitude of the current, we can conclude that negligible current flows in the high-doped body of the device with the halo, which is not the case in the device without it. Because of a higher Gummel number in the halo region than in the body, defined by the higher doping density, the electron injection into the halo region is suppressed, and the high-doped body becomes relatively inactive for the BJT transport current; whereas the injection into the low-doped region is active in both cases. The injection (across the sidewall of the source-body junction) into the low-doped body is controlled by the V_{GS}-dependent majority carrier density, which is lower than the doping level in the region even when the halo wraps around the source. Thus the bipolar current



(a)



(b)



(c)

Fig. 4.2 PISCES-simulated electron transport current in BTS NFD/SOI MOSFET: (a) the simulated structural approximation for BTS device; (b) the electron current in the device with halo; (c) the electron current in the device without halo.

gain, β, is reduced mainly because of the reduction of the effective emitter area adjacent to the high-doped body (bas) region. This reduction implies higher BV [Cho91b] in the device with the halo structure.

However, implicit in the simulation results is the reduction of the pinched resistance due to halo region from the body contact to the body region, modulated by the doping density in p^+ region in Fig. 4.2.(a). In addition to the reduced β, the halo results in a smaller R_B because of the higher doping density, which leads to a reduced ohmic voltage drop across it for a given V_{DS}-dependent impact-ionization current. Therefore, at the same V_{DS}, this voltage drop in halo device may not be sufficient to turn on the parasitic BJT, which also results in increased BV.

To examine the relative effect on the breakdown voltage due to R_B and β, we turn to SOISPICE simulations. Figure 4.3 shows the SOISPICE-simulated breakdown characteristics with and without halo. The model parameters for the simulated halo device are extracted based on the halo-device structure in Fig. 4.2.(a). Then the effects of R_B and β are investigated by varying them respectively and independently for no-halo simulations in Fig. 4.3. R_B is increased for one no-halo simulation by a factor of 10 (the ratio of concentration variation in PISCES simulation). The β variation is effected in the other no-halo SOISPICE simulation by setting THALO = 0; when THALO = 0, J_{N0} for a device without halo increases according to (2.64) from the value corresponding to the halo device, effecting an increased β if other parameters remain fixed. As a result, a discernible shift in breakdown voltage

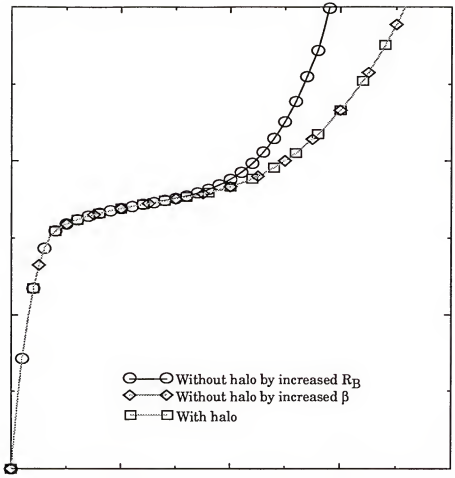


Fig. 4.3 SOISPICE-simulated breakdown characteristics of a typical n-channel BTS NFD/SOI MOSFET, showing the effects of R_B and β on the breakdown characteristics.

can be seen in Fig. 4.3 only when R_B is increased, suggesting that the breakdown voltage increase in a device with halo is primarily due to R_B variation. Therefore, in the analyses and discussions of the following sections, we will concentrate on the effect of R_B, assuming a constant β which is not unreasonable near the breakdown.

4.3 Derivation of Breakdown Conditions

Figure 4.4 shows a typical (SOISPICE-simulated) breakdown characteristics of an n-channel BTS-NFD/SOI MOSFET with finite (lumped) R_B. The holding voltage, V_H, and the snapback voltage, V_S, are distinctively labeled in the figure. At all bias conditions, the drain current is defined by impact-ionization multiplication of the parasitic BJT current ($I_T = \beta I_R$) as well as of the channel current (I_{CH}):

$$\begin{aligned} I_D &= M(I_{CH} + I_T) \\ &= \frac{MI_{CH}}{1-\beta(M-1)I_R/I_{Gi}} \end{aligned} \tag{4.1}$$

where $I_{Gi} = (M\text{-}1)I_D/M$ is the impact-ionization current and $I_R \approx WJ_{S0}\exp(V_{BS}/V_{TH})$ is the recombination current in the source [Cho91b]; V_{TH} is the thermal voltage, and V_{BS} is the induced body-source junction forward bias defined by R_B and the division of generated hole flow:

$$V_{BS} = R_B(I_{Gi} - I_R)\,. \tag{4.2}$$

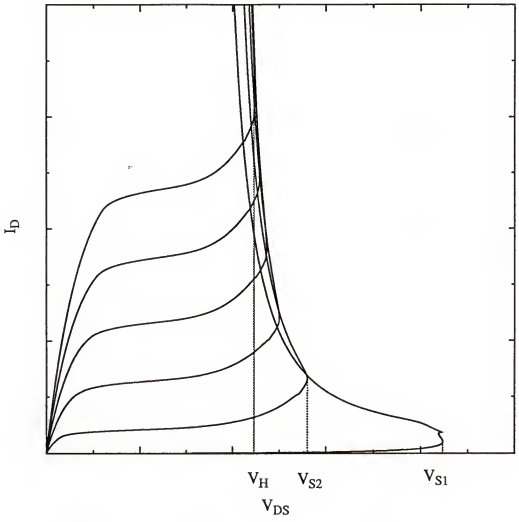


Fig. 4.4 Typical breakdown characteristics of an n-channel BTS-NFD/ SOI MOSFET.

Note that $I_{Gi} > I_R$ except in the floating-body case ($R_B = \infty$), for which $I_{Gi} = I_R$.

The holding-voltage condition can be defined where I_D and dI_D/dV_{DS} are infinity, and the snapback condition where only the derivative is infinity. Assuming constant β, which is not unreasonable near the breakdown, we derive from (4.1)

$$\frac{dI_D}{dV_{DS}} = \frac{f'(1-g) + fg'}{(1-g)^2} \tag{4.3}$$

where $f = MI_{CH}$ and $g = \beta(M-1)I_R/I_{Gi}$; the primes denote derivatives with respect to V_{DS}. By examining (4.1) and (4.3) then, we conclude that the holding-voltage condition is achieved when $g = 1$:

$$(M-1)\beta\frac{I_R}{I_{Gi}} = 1 \qquad \text{at } V_{DS} = V_H. \tag{4.4}$$

For the snapback condition, g must not be equal to 1 in order for the drain current to be finite. Since f, f′, and g are finite quantities in (4.3), snapback occurs when g′ becomes infinity. To express the condition, we rewrite $g = I_T/(I_{CH} + I_T)$, and take the derivative with respect to V_{DS}. Note that prior to snapback, $I_{CH} \gg I_T$ and $g \approx 0$. Thus we get

$$g' = \frac{I_T'(1-g) - gI_{CH}'}{I_{CH} + I_T} \approx \frac{I_T'}{I_{CH}} \tag{4.5}$$

where, consistent with the constant-β assumption,

$$I_T' = \frac{I_T}{V_T} V_{BS}' . \tag{4.6}$$

To express V_{BS}', we rewrite (4.2) as

$$V_{BS} = R_B \{ (M-1) I_{CH} + [\beta (M-1) - 1] I_R \} \quad , \tag{4.7}$$

and differentiate to obtain

$$V_{BS}' = [(M-1) I_{CH} R_B]' + R_B \beta (M-1)' I_R + I_R' R_B [\beta (M-1) - 1] . \tag{4.8}$$

Substituting (4.8) into (4.6) and rearranging terms, we have

$$I_T' = \frac{I_T}{V_T} \left(\frac{[(M-1) I_{CH} R_B]' + R_B \beta (M-1)' I_R}{1 - \frac{I_R R_B}{V_T} [\beta (M-1) - 1]} \right) . \tag{4.9}$$

When the denominator of (4.9) becomes zero, I_T' becomes infinity and thus so does g′ in (4.5). Hence, the snapback condition is achieved:

$$\frac{I_R R_B}{V_T} [\beta (M-1) - 1] = 1 \qquad \text{at } V_{DS} = V_S . \tag{4.10}$$

4.4 Discussion

With a finite R_B, a fraction of the generated holes flow out the body contact, creating a voltage drop across R_B which is the forward bias V_{BS} on

the source-body junction in (4.2). Thus the ratio I_R/I_{Gi}, which defines the holding voltage in (4.4), is less than unity. This is in contrast to the floating-body case ($R_B = \infty$) where I_R/I_{Gi} equals unity. The SOISPICE simulation results in Fig. 4.5 for finite R_B show that typically the ratio is significantly less than unity even when the device is in the breakdown regime; the current through R_B is never negligible when compared to I_R. The lower the R_B is, the smaller the ratio is and the greater (M-1)β must be at $V_{DS} = V_H$. Only for the floating-body case is the ratio unity and (M-1)β = 1. In the ideal case of zero body resistance, the ratio is zero and the device never reaches the holding-voltage condition. This dependence on R_B underlies the increase of V_H afforded by the BTS device as compared to the floating-body counterpart.

If we assume that $V_{BS} \approx 1.0$ V at $V_{DS} = V_H$, the holding-voltage condition (4.4) can be approximately characterized, using (4.2), as

$$(M-1)\,\beta \approx \left(1 + \frac{1.7 \times 10^{-17}}{W J_{R0} R_B}\right) \qquad \text{at } V_{DS} = V_H, \quad (4.11)$$

with R_B given in Ω and WJ_{R0} in A. Note that W is the effective width of the device. As (4.11) clearly shows, the increase of (M-1) due to finite R_B, which implies the increase in V_H, is inversely proportional to R_B. However as revealed in Fig. 4.6, the improvement is not significant for a typical J_{R0} unless R_B is extremely low. Figure 4.6 shows holding-voltage improvements versus R_B derived from the SOISPICE simulations. The simulated NFD/SOI MOSFET is a typical short-channel device: L = 0.5 μm, W = 2.4 μm, t_f = 300 nm,

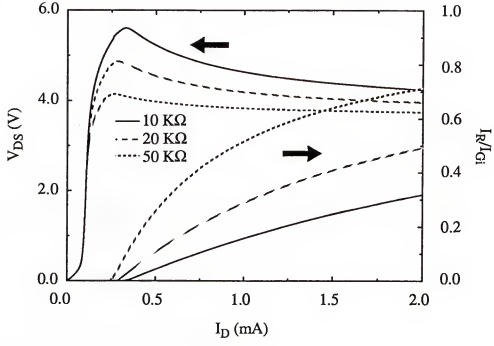


Fig. 4.5 SOISPICE-predicted I_R/I_{Gi} ratio for different values of R_B, and corresponding I_D - V_{DS} curves; V_{GfS} = 3 V. The NFD device simulated is typical: W/L = 2.4 μm/0.5 μm, t_b = 300 nm, and N_{BL} = 1.2 x $10^{17}/cm^3$, but without LDD/LDS.

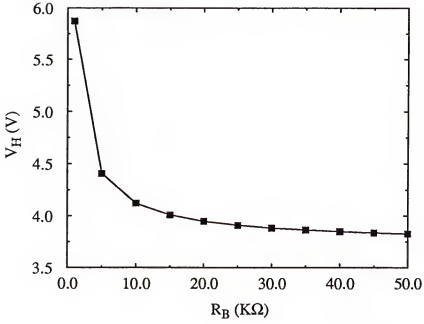


Fig. 4.6 SOISPICE-predicted holding voltage versus body resistance for a typical NFD/SOI MOSFET (L = 0.5 μm), but without LDD/LDS.

N_{BL} (average channel doping density) = 1.2 x 10^{17} cm^{-3}, but no LDD/LDS. From the physical parameters, J_{R0} is calculated in SOISPICE to be ~4 x 10^{-21} A/μm. For this device, R_B « 10 kΩ is needed to achieve a significant increase in V_H. This result translates to the requirement that a BTS structure must reduce the specific body resistance to R_BW « 24 kΩ-μm in order to yield a substantive increase in V_H. This conclusion applies to LDD devices as well. Typically however, BTS structures, in which the strap width (W_{straps}) is limited relative to W, do not reduce R_BW this much [Hwa91]. Since $R_B \propto W$, the required inequality inferred above defines an upper limit on W. Therefore current drive ($\propto W$) requirements will dictate more straps for an effective BTS, and hence a total device width ($W + W_{straps}$) that is prohibitively larger than W. For example, if the body sheet resistivity were 12 kΩ/□, which is representative of NFD/SOI MOSFETs [Hwa91], then the requirement on specific body resistance implies W« ~2 μm, which in turn necessitates a multiple-strap design in which W must be comparable to W_{straps}.

The snapback condition described by (4.10) is dependent not only on R_B but also on I_{CH}. Until V_{DS} reaches V_S, most of the impact-ionization generation current flows through R_B. Therefore from (4.2), $V_{BS} \simeq (M-1)\, I_{CH} R_B$, and this induced voltage defines I_R. Thus (4.10) can be rewritten as

$$\frac{WJ_{R0}R_B}{V_T}[\beta(M-1)-1] \simeq \exp\left(\frac{I_{CH}R_B(M-1)}{V_T}\right) \text{ at } V_{DS} = V_S; \quad (4.12)$$

(M-1) defined by (4.12) implies the snapback voltage. Qualitatively, V_S is the drain bias at which the BJT transport current I_T becomes significant in comparison to the channel current I_{CH}. (Remember that $I_T' = \infty$ at $V_{DS} = V_S$.) As I_{CH} increases for a finite R_B, V_S decreases as characterized in (4.12). This is logical because the larger the I_{CH} is, the smaller the (M-1), and hence the lower V_{DS}, need be to turn the BJT on. For continued increase of I_{CH}, V_S decreases until it ultimately equals V_H characterized in (4.11). In this case, snapback characteristics do not appear as indicated in Fig. 4.4. As R_B increases for a finite and low channel current, V_S also approaches V_H, both of which are decreasing to a value given by the condition $(M-1)\beta = 1$ in the limit of infinite R_B (floating-body case). Only when V_{DS} reaches a snapback voltage significantly larger than V_H does the I_D-V_{DS} curve show the negative-resistance characteristic. In Fig. 4.7, we compare (M-1) predicted by (4.12) with that derived from the SOISPICE simulation results at $V_{DS} = V_S$ for varying R_B. There is a very good agreement.

In the negative-resistance region, a sufficiently high voltage across the body resistance is maintained such that the BJT is in a constant-current gain mode. Indeed, in SOISPICE, β is constant near the breakdown, and still the simulations show negative-resistance curves. Thus we conclude that the negative-resistance characteristic of the BTS device results because the loop gain reflected by (4.1) increases due to increasing I_R/I_{Gi} as the holding-voltage

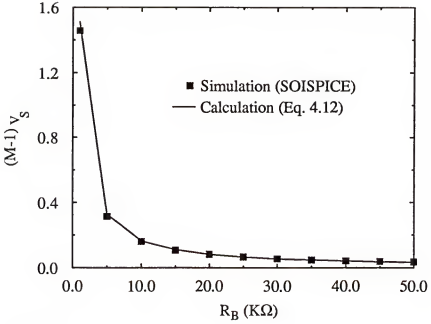


Fig. 4.7 Comparison of calculated and simulated (M-1) at $V_{DS} = V_S$.

condition is approached; it is not associated with an increasing β, which underlies snapback in common-emitter BJT breakdown characteristics.

4.5 Conclusions

Breakdown characteristics of BTS NFD/SOI MOSFET with halo were investigated with the aid of 2-D device simulator and SOISPICE. The simulations revealed that BV increase in halo device is mainly due to the reduction of the parasitic R_B, which is modeled as a lumped resistance in the SOISPICE simulations. Simple but physical characterizations of the holding voltage and the snapback voltage of the (BTS-NFD) SOI MOSFET have been derived with the aid of SOISPICE simulations. The characterizations showed dependences on the body resistance, the carrier recombination property of the body-source junction, and the channel current. The holding voltage increases as R_B decreases, but for submicron devices the improvement is insignificant unless the specific body resistance R_BW is much less than a critical value. This critical value is much less than values yielded by typical BTS structures in which the strap width is limited relative to W. The negative-resistance snapback, corresponding to $V_S > V_H$, is an inherent effect of finite body resistance, and is not due to variation of the parasitic BJT β. The snapback voltage is dependent on the channel current as well as R_B; V_S decreases as R_B and I_{CH} increase. The physical insight afforded by the analysis and (4.11) and (4.12)

should be useful in the optimal design of BTS-NFD/ SOI MOSFETs, and indeed in assessing the viability of BTS structures.

CHAPTER FIVE
ASSESSMENT OF THE FLOATING-BODY EFFECTS IN NFD/SOI CMOS DEVICES AND CIRCUITS

5.1 Introduction

In the previous chapter, we examined the efficacy of body ties in removing or ameliorating unwanted floating-body effects, especially the premature breakdown. As discussed and evidenced, however, unless the parasitic body resistance, R_B, is effected below a very small value, the utility of body ties is questionable; the increase of device area to achieve small enough R_B is prohibitive in cost, which undermines the viability of the body-tie technology. Therefore, for a viable NFD/SOI technology, the floating-body configuration seems to be the only option.

As revealed in Chapter 3, the floating body causes numerous so-called floating-body effects in device characteristics even at V_{DS} below breakdown. The DC effects are schematically shown in Fig. 5.1. The floating-body voltage, V_{BS}, is induced by free carrier charging from various generation mechanisms which is balanced by the recombination in the device. Except for the BJT-induced breakdown, the DC floating-body effects are due to the sensitivity of the threshold voltage of the NFD device to the floating-body voltage, V_{BS}. In a viable NFD/SOI CMOS technology, all of the floating-body effects

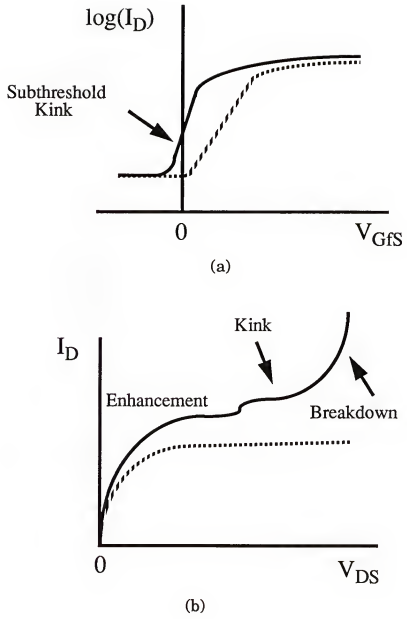


Fig. 5.1 Schematic characteristics of a floating-body n-channel NFD/SOI MOSFET (solid lines) in contrast to tied-body characteristics (broken lines): (a) subthreshold characteristics, (b) I_D-V_{DS} characteristics.

must be recognized, and either inhibited or exploited pragmatically. Obviously, the BJT-induced breakdown is detrimental and must be controlled; the potential benefit afforded by the supplemental BJT current drive in digital applications is small and is preempted by the associated loss of gate control [Fos93a]. However, the effects due to V_{BS}-induced threshold reduction could possibly be exploited pragmatically to improve performance. Since the floating-body effects in Fig. 5.1 (premature breakdown, saturation-region and subthreshold kinks, and low-V_{DS} current enhancement) are modeled physically well by SOISPICE, the simulations using it can thus give physical insight regarding the effects and their influences on NFD/SOI circuits.

However, the aforementioned floating-body effects should not translate into circuit-level merits or demerits in a straightforward manner. The fluctuation of the floating-body charge conditions in dynamic circuit environments could affect the generation-recombination balance, and thus could control the transient floating-body voltage in a more complex manner. Furthermore, in an extreme situation, a transient BJT current can be induced [Pel95]. The reliable device/circuit simulation capability afforded by SOISPICE with our physical and compact NFD SOI MOSFET model, which includes the unique accountings of the body-charge dynamics, can be used to check this exploitation of the floating-body effects in more comprehensive aspects.

In this chapter, we examine the floating-body effects in NFD/SOI CMOS devices and circuits to provide the insights for an optimal device/circuit

design. The conclusions from the analyses also lead to discussions of the viability of the NFD/SOI technology as a serious contender to the bulk-silicon technology.

5.2 DC Floating-Body Effects

All of the floating-body effects under DC conditions in Fig. 5.1 are due to excess majority-carrier charge in the body, injected via carrier generation, which can be removed only via recombination. This charging effect is accounted for in SOISPICE by the nodal equation governing B' node in Fig. 3.1. Finite carrier lifetimes thus define a forward bias V_{BS} on the body-source junction which drives the mentioned effects. The premature breakdown results when V_{BS} activates the parasitic BJT and when V_{DS} is high enough that the BJT current drives significant impact ionization, which in turn injects more charge in the body and causes the process to become regenerative [Fos93a]. Thus, there is a consensus to inhibit this effect by avoiding such a high V_{DS} under normal device operations.

The low-V_{DS} effects are due to the sensitivity of the NFD threshold voltage to V_{BS} (which is negligible in the FD device). The kink effects are driven by impact ionization as well, but at values of V_{DS} too low to cause regeneration. In strong inversion, the kink in saturation current results from V_{BS}-dependent threshold reduction, as V_{BS} is limited to a "diode drop"; in weak inversion, the kink is reflected as an abnormally high slope, or low gate voltage swing S, defined by the increasing V_{BS} with V_{GfS} [Fos87]. The low-V_{DS}

current enhancement, which occurs prior to the kink, is not well known. It is driven by thermal carrier generation. In this section, we analyze these effects respectively using SOISPICE device/circuit simulations.

5.2.1 Kink Effect

Exploitation of the kink effects is risky since they are driven by impact ionization, which also drives the parasitic BJT and the ultimate loss of gate control. It is obvious from Figs. 3.6(b) and 3.7(a) that attempting to design for a supply voltage above the saturation-current kink but below breakdown is nonsensical. However it is not so obvious that designing to exploit the subthreshold kink (Fig. 3.7(b)) is impractical. Very low S, much lower than the ideal 60 mV, is possible [Fos87]. If such a low S could be achieved in a viable design, then possibly the threshold voltage could be reduced to get higher current drive and faster speeds, without threatening off-state ($V_{GfS} = 0$) current and standby power restrictions [Sha93].

Insight from our model suggests that $S < 60$ mV can be achieved by controlling the $I_R(V_{BS})$ dependence in (2.66), while still inhibiting the BJT. From (2.47) and (2.48), we can approximate the gate voltage swing as

$$S = \left(\frac{d\log I_D}{dV_{GfS}}\right)^{-1} \cong \frac{S_0}{\left(1+\alpha\frac{\partial V_{BS}}{\partial V_{GfS}}\right)} \tag{5.1}$$

where $S_0 = V_{TH}(1+\alpha)\ln(10)$, which is the normal value of S when $V_{BS} = 0$. (Note that $\alpha > 0$ defines $S_0 > 60$ mV for the NFD/SOI MOSFET.) The reduction

of S in the floating-body device is characterized by the derivative in the denominator of (5.1), which is defined by I_R and the impact-ionization current I_{Gi}. Since $I_R = I_{Gi} \cong (M-1)I_{CH}$,

$$V_{BS} = mV_{TH}\ln\left[\frac{(M-1)I_{CH}}{WJ_{R0}} + 1\right]. \tag{5.2}$$

Differentiating (5.2) with respect to V_{GfS} yields

$$\frac{dV_{BS}}{dV_{GfS}} \cong mV_{TH}\ln(10)\frac{d\log I_{CH}}{dV_{GfS}}, \tag{5.3}$$

if we assume $I_{Gi} >> WJ_{R0}$. Thus (5.1) can be rearranged using (5.3) as

$$S = S_0 - m\alpha V_{TH}\ln(10)\,. \tag{5.4}$$

Note from (5.4) that when the nonideality factor m for I_R is unity, S = $V_{TH}\ln(10)$ = 60 mV. Thus, only when m > 1 can S be less than the ideal value. Such is the case in typical NFD/SOI MOSFETs as evidenced in Fig. 3.7(b). In essence, a higher m defines a higher V_{BS} for a given I_{Gi}, and hence a more significant subthreshold kink effect. This dependence on m is illustrated in Fig. 5.2 where SOISPICE-simulated subthreshold characteristics for varying m are shown. Note that for m = 2 which is a physical upper limit, S is extremely low but the off-state current is very high. In several measured characteristics [Fos87] [Dav86], S less than what is dictated by m=2 have been reported, which is as low as 10mV. In such cases, the BJT current is significant in the subthreshold characteristics, and the device is close to a latch condition (a. la.

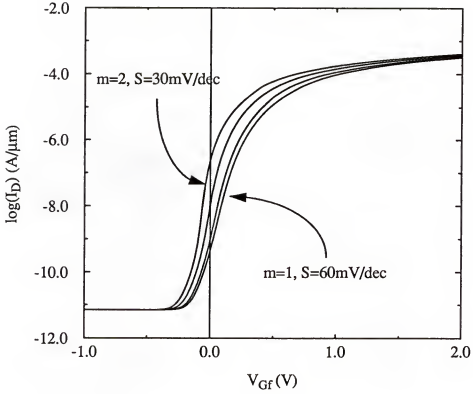


Fig. 5.2 The predicted subthreshold characteristics as the non-ideality factor, m varies. The device of L_{eff} = 0.2μm is simulated, biased at V_{DS}=2V. No LDD/LDS is included.

BJT-induced breakdown) [Fos93a]. Operating a device in such extremity is not desirable. Therefore we will neglect such a case in the following considerations.

The high off-state current is a problem, but does the abnormally low S actually yield a speed benefit? SOISPICE simulations suggest otherwise. Figure 5.3 shows SOISPICE-predicted propagation delays versus supply voltage taken from simulations of CMOS inverter chains using different values of m in the I_R model. The delays predicted when the impact ionization was turned off are included for comparison. Interestingly, the subthreshold kink, irrespective of m, results in no speed improvement. Neither could we observe speed improvement due to kink in Fig. 5.4 where SOISPICE-predicted propagation delays are also shown using the same inverter chain in Fig. 5.3 but with different values of capacitive loads at each stage. These results can be attributed to the dynamic nature of the inverter circuit. Simply stated, the added drive current afforded by the subthreshold kink is available only for high V_{DS}, but V_{DS} is not high during most of the switching cycle. For example, when the n-channel MOSFET is switched on and begins to discharge the load, V_{DS} drops, and the added drive current is lost long before the switching is completed. In addition, the effect of threshold voltage reduction due to kink is relatively greater at lower V_{GfS}, if it is. However, at low V_{DD} which renders lower V_{GfS}, V_{DS} is not high enough to effectuate the subthreshold kink. We conclude then that the kink is not beneficial, and in fact is detrimental with

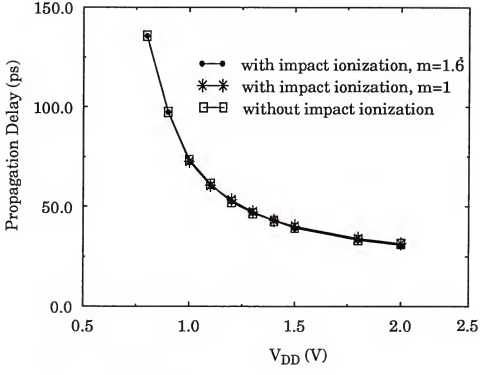


Fig. 5.3 SOISPICE-predicted propagation delay with (for different m) and without kink versus the supply voltage, V_{DD}.

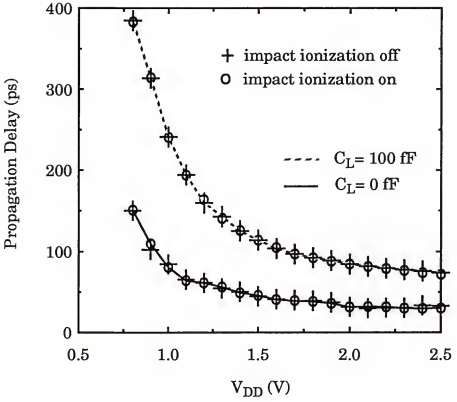


Fig. 5.4 SOISPICE-predicted propagation delay with and without kink versus the supply voltage, V_{DD}, for different capacitive loads.

regard to off-state current. It must be suppressed, which is compatible with inhibiting the BJT-induced loss of gate control.

5.2.2 Low-V_{DS} Current Enhancement Effect

Next we consider the low-V_{DS} current enhancement evident as a floating-body effect in Fig. 3.6. This effect is driven by thermal generation current (I_{Gt}). For DC conditions at low V_{DS} (where impact ionization is negligible), $I_R = I_{Gt}$ defines

$$V_{BS} = mV_{TH}\ln\left(1 + \frac{I_{Gt}}{WJ_{R0}}\right). \tag{5.5}$$

Depending on the ratio I_{Gt}/WJ_{R0}, this body bias can reduce the device threshold voltage and thereby enhance the current. Typically, as reflected by Fig. 3.6, V_{BS} in (5.5) can be a few tenths of a volt, which can produce a significant current enhancement. The SOISPICE-predicted I_D-V_{DS} characteristics in Fig. 5.5 confirm this significance. The simulations were done with the same set of (realistic) model parameters (with L = 0.2 μm) used in Fig. 3.7, but with the impact ionization turned off to emphasize the I_{Gt} drive. The difference in saturation currents for the tied-body and floating-body configurations is substantial. The current enhancement is about 50% at $V_{GfS} = V_{DS} = 1.0$V and about 20% at $V_{GfS} = V_{DS} = 2.0$V, suggesting exploitation of this effect for low-voltage CMOS applications. Indeed, V_{BS} governed by (5.5) and the concomitant V_T-lowering is practically constant, and its effect becomes relatively greater at

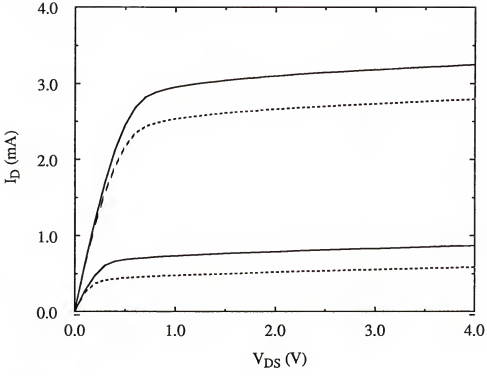


Fig. 5.5 SOISPICE-predicted I_D-V_{DS} characteristics of an n-channel NFD/SOI MOSFET for floating-body (solid line) and tied-body (broken line) connections with impact ionization off (V_{GfS}=1V, 2V, L_{eff}=0.2μm), showing the significance of I_{Gr}-induced V_T lowering.

lower V_{GfS} (V_{DD}). Note that the enhancement could be partly due to an intrinsic $V_{BS} < 0$ in the tied-body device, which increases the threshold voltage, because of internal ohmic drops in the source region. Nonetheless, the advantage of the floating body in this regard is palpable.

In transient operation of the device, this current enhancement will prevail; a "latent" $V_{BS} > 0$, driven by I_{Gt} as described in (5.5), will be sustained for all finite V_{DS}, even though the body charging current (dQ_B/dt in Fig. 3.1) will tend to modulate it. Since the switching frequencies are high, as we will discuss later, the modulation will not generally be significant during switching. For this same reason, the different body charge dynamics in different devices in the same circuit, which if significant would render exploitation of this floating-body effect impractical, should be inconsequential. This insight is confirmed by the SOISPICE-predicted propagation delays versus supply voltage (V_{DD}) plotted in Fig. 5.6. The delays were derived from simulations of CMOS inverter-chain circuits using the devices represented in Figs. 3.7 and 5.5, for tied-body and floating-body configurations. A significant speed improvement due to the latent V_{BS} in the floating-body device is predicted. As V_{DD} is scaled down, the delay increases due to smaller current drive, but the relative magnitude of the speed benefit due to the latent V_{BS} increases. Thus this particular floating-body effect appears to be quite advantageous for low-voltage CMOS applications, and its benefit will be more pronounced the lower V_{DD} is and the higher the (capacitive) loads are. The effect implies a significant advantage of SOI over bulk CMOS in low-voltage applications. However,

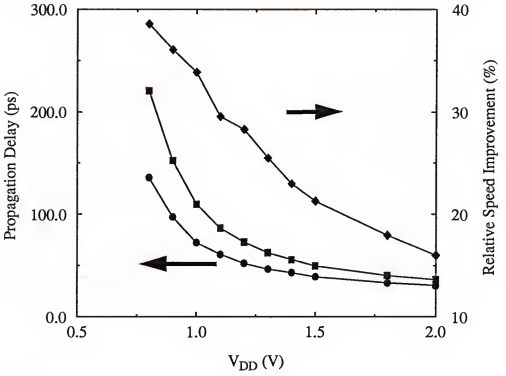
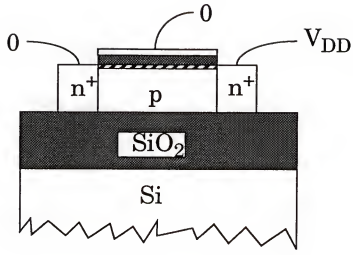


Fig. 5.6 The SOISPICE-predicted propagation delay vs. supply voltage, V_{DD} for floating (●) and tied (■) bodies. The delays are derived from unloaded three-stage NFD/SOI CMOS inverter-chain simulations. Also shown is the predicted relative speed enhancement (♦) afforded by the floating body.

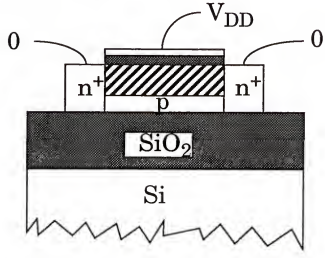
the speed improvement would seem to imply a common design tradeoff because it results from effective threshold lowering which would also increase off-state current and hence static power. Fortunately, the transient circuit simulations suggest though that this tradeoff is not needed; they show that the benefit is not undermined by increased off-state current because of a unique dynamic floating-body effect in the NFD/SOI MOSFET, which will be discussed in the next section.

5.3 Transient Gate-Induced Floating-Body Effects

DC floating-body effects discussed in the previous section are compounded in dynamic circuit operations by the unique floating-body charge dynamics. Figure 5.7 shows the steady-state body charge conditions under two frequently encountered bias sets in digital circuits, i.e. off and on states. As the NFD MOSFET is switched back and forth between these two states, the body of the device tends to recover from one to the other steady-state charge condition. Whereas this recovery is very fast (subject to dielectric relaxation time) in bulk devices (or body-tied devices with sufficiently small body resistance), it is relatively slow (subject to recombination and generation lifetimes) in floating-body devices. Consequently, the floating-body charging/discharging lags the terminal voltages in contemporary high frequency applications. Thus, a transient floating-body voltage, $V_{BS}(t)$, which could be positive or negative, develops during switching, and it can be sustained even after the terminal voltages reach a steady state while the floating-body charge condition is being



(a)

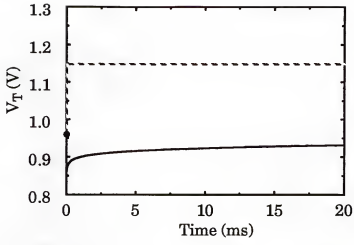


(b)

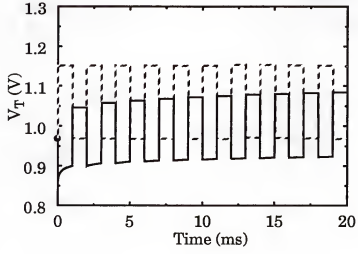
Fig. 5.7 Schematic cross-sectional view of a n-channel NFD/SOI MOSFET showing steady-state floating-body charge conditions under two frequently encountered bias conditions in digital circuits. The hatched area represents relative degree of depletion under the gate; (a) off state, (b) on state.

fully recovered. For example, when the device is switched from off to on (a to b in Fig. 5.7), the body-charge condition in Fig. 5.7(b) is reached by removing the excess majority carriers from Fig. 5.7(a) through recombination, which causes a positive $V_{BS}(t)$. In the other direction (b to a in Fig. 5.7), $V_{BS}(t)$ becomes negative as majority carriers are generated to neutralize the depletion charge, and produce the state represented in Fig. 5.7(a). This gate-induced body-charge dynamics is very important in NFD/SOI device and circuit applications because $V_{BS}(t)$ defines a dynamic threshold voltage, $V_T(t)$, and a transient BJT current, $I_T(t)$, as well.

The $V_T(t)$ problem is exemplified in Fig. 5.8 where SOISPICE-predicted threshold voltage of a typical short-channel (0.2 μm) device are shown versus time for the gate switching from on and off steady states. The variation of V_T in time reflects $V_{BS}(t)$ implicitly. The initially identical threshold voltages of the device implies only that the device begins from a steady state ($V_{BS}(0) = 0$) in both cases. However, the initial body-charge conditions are different, and the initial charge dictates the direction of the threshold variation in Fig. 5.8(a) upon the subsequent switching. When the device is initially off and then switched on, a positive $V_{BS}(t)$ whose value is the greatest at $t=0^+$ and decays as time goes on, develops as the excess free carriers recombine across the junction. The initial threshold voltage reduction and subsequent increase in time implies this. In the opposite case where the device is initially on and switched off, the depleted body forces $V_{BS}(t)$ to be negative after switching as the body is in need of free carriers to replenish it. Therefore the threshold volt-



(a)



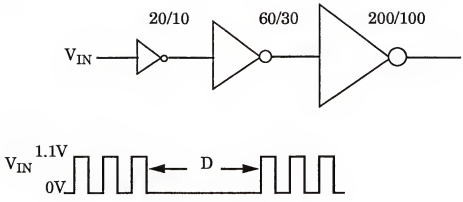
(b)

Fig. 5.8 SOISPICE-simulated transient threshold voltage of a floating-body n-channel NFD/SOI MOSFET (L = 0.2 μm), corresponding to 0 V (solid line) and 1.5 V (broken line) as initial V_{GfS}'s. The applied gate voltages are a 1.5V-step function in (a) and a 500kHz pulse in (b)

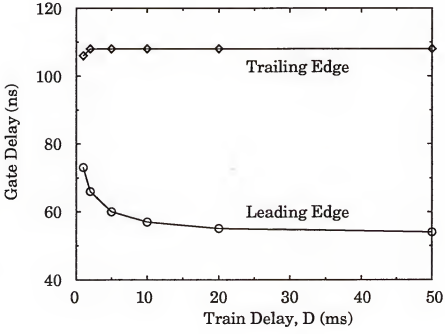
age initially increases, and is sustained even longer due to limited generation capability of the device, even though it will eventually come back to the DC value in time. The time constants of the recovery mechanisms, as implied in Fig. 5.8(a), are very long; they are about 10 ms for the charging (carrier-replenishing) recovery and about 1 μs for the discharging (carrier-removing) recovery in a typical contemporary scaled NFD/SOI MOSFET. The dependence of the recovery time on the switching direction is due to the fact that in one case it is dependent on the forward-bias (recombination) characteristics, and in the other case on the reverse-bias (generation) characteristics of the body-source junction. Furthermore, if the device is switched with a finite frequency as in Fig. 5.8(b), this charge dynamics become more complicated because of the counteracting tendencies of recombination and generation, which now alternate. Therefore, at a given point in time the device's threshold voltage is dependent on the recent switching; it is unknown unless this history is known.

5.3.1 Triple-Pulse Gate Delay

Propagation delay of an inverter circuit will be influenced by the hysteretic $V_T(t)$ due to the transient floating-body charging effect described above. To show this effect, measurements were taken on a three-stage CMOS inverter buffer; results are shown in Fig. 5.9. The measurements were done with a triple-pulse train [Hou95], where the interval between trains is varied to allow different floating-body charge conditions. The transistors in the circuit were standard floating-body devices in which the bodies are all



(a)



(b)

Fig. 5.9 Triple-pulse experiment; (a) schematic diagram of the NFD/SOI CMOS inverter with the input-pulse definition; (b) experimentally extracted gate delay of the first leading and trailing edges versus train delay. L=1μm. The numbers above the inverter schematic represents W_P/W_Nin μm for each stage.

individually isolated and left floating. The gate delays are taken on three leading edges (input going low to high) and three trailing edges (input high to low) of the triple pulses. The gate delays were taken after sufficient numbers of pulse trains were applied to obviate the unknown history and define a definitive starting steady state.

The delays plotted in Fig. 5.9(b) clearly support the floating-body charging effect depending on the interval between trains, and on the leading or trailing aspect of the switching signal. The leading-edge delays are shorter than the trailing-edge ones because the leading-edge switching is driven by the devices that stayed longer in the off state, and thus experience reduced $V_T(t)$ during transitions. The leading-edge delay increases as the period between trains diminishes; the floating-body charge in the off transistors is less recovered during the diminished interval, and $V_{BS}(t) > 0$ in subsequent switching is smaller. The trailing-edge delay however only slightly decreases with decreasing period between the pulses; during the reduced period, the on-transistors may not completely discharge the body charge restored in the devices previously while the pulse voltage is high.

The transients of the triple-pulse experiment involve significantly different time constants. The time-step control algorithm in SPICE is not adequate to handle the large time-constant variation in these transients. Therefore, two-part simulations were done to separate the time constants; the pulse part and the delay part are simulated alternately. The final transient voltage solution of one part is passed to the other part as the initial condition,

i.e., via the ICs on the device line of the circuit file. The two-part simulations are continued until the final solution of the delay part has converged to the initial conditions of the previous pulse-part simulation; the two-part simulated circuit has then reached a steady state, and the gate delays are taken from the subsequent pulse-part simulation.

SOISPICE simulations of this transient experiment give reasonable results, supporting the essence of the charge modeling. The predictions of the triple-pulse delay simulation are shown in Fig. 5.10, where the model parameters used in the buffer simulations were evaluated from DC I-V measurement data and structural information on similar devices provided by [Hou95]. Note that using the parameters extracted via the scheme in Sec. 3.3 will result in discrepancies in the high-power region of I-V data that could lead to an overestimation of current drive, hence smaller gate delay. The trend in predicted delays follows that in experiment, but with a constant shift in both leading and trailing edge delays; the predicted delays are about 20ns less than the measured ones.

The discrepancy can be partly due to circuit parasitics unaccounted for in the simulations, but also possibly due to deficiencies in the present SOI technology, such as nonuniformity of the wafer, e.g., inhomogeneous defect densities, and the implied uncertainty of the measured DC characteristics. Such inhomogeneities, as well as self-heating and floating-body dynamics, suggest that the actual characteristics of the devices in the buffer circuit can be different from the test-device characteristics to which the model parame-

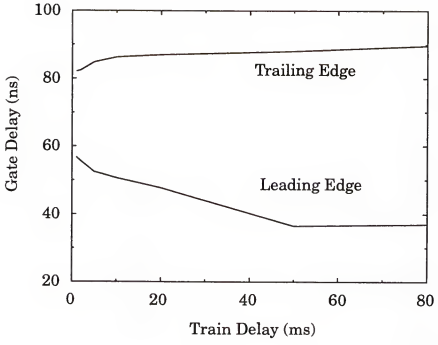


Fig. 5.10 SOISPICE-predicted gate delays of leading and trailing input transitions for the CMOS NFD/SOI inverter buffer in Fig. 5.9.

ters are tuned. Such mismatches are reflected more readily in low-voltage circuit simulation results.

5.3.2 Implications on Pulsed I-V Measurement

Self-heating in the SOI MOSFET, be it FD or NFD, affects DC measurement data especially in high-power regions. Heat generated by the IV power during measurement of SOI MOSFETs cannot be dissipated off as easily as in bulk devices because of thick back oxide with low heat conductivity. Since the self-heating is mainly a DC effect [Su94], model parameters for digital circuit simulations must be evaluated without it. Our model-parameter extraction scheme described in Sec. 3.3 can exclude the self-heating effect if based on the low-power regime of I-V data. However, we can infer that the extracted model parameters will result in relevant predictions in circuit simulations in spite of the discrepancy due to insignificance of self-heating in digital circuit operations.

Conventional parameter extraction/optimization cannot be done without I-V data exclusive of self-heating, for it is based on curve fitting without much physical implications. Recently, a pulsed I-V measurement method [Jen94] was proposed to avoid self-heating. However, although unacknowledged, this method can be undermined by dynamic floating-body charging effect. To efficiently avoid the self-heating, the pulse intervals between samplings should be as large as possible while the pulse width should be minimal. These requirements on pulses tend to accentuate the dynamic charging effects. As exemplified in Fig. 5.11, the SOISPICE-predicted transient drain

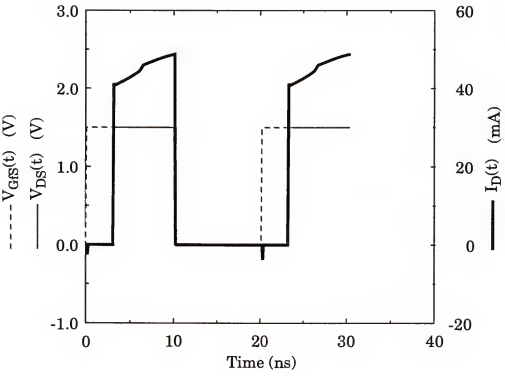


Fig. 5.11 SOISPICE simulation of a pulsed I-V measurement on a floating-body n-channel NFD/SOI MOSFET.

current for gate and drain voltage pulses with widths (nanoseconds) comparable to those used in [Jen94] show that the current might not reach its static level, and hence the parameter values extracted based on such data could be invalid. The delay of drain current in Fig. 5.11 is the time required for generated (e.g., via impact ionization for high V_{DS}) carriers to charge the body. When V_{DS} is low, however, the pulse measurement will result in an overestimation of the current due to the transient current overshoot effect [Lim84]. The time to reach general steady state depends on the device structural and electrical properties, and hence proper voltage pulse widths are not readily apparent.

5.3.3 Transient-Induced Current Undershoot

The threshold voltage transients in Fig. 5.8 imply transient current undershoot or overshoot, which could be beneficial in circuit performance. As indicated by the SOISPICE-predicted transient currents in the CMOS inverter chain plotted in Fig. 5.12 for floating and tied bodies, the dynamic off-state current is effectively suppressed when the body is floating. The suppression occurs because the MOSFET body is partially depleted of majority carriers in the on state, and the depletion charge in the off state forces $V_{BS}(t) < 0$ as the carriers are recovered, as in Fig. 5.12(b). Whereas the time constant associated with this recovery in the tied-body device (i.e., the dielectric relaxation time) is negligibly short and thus V_{BS} becomes zero almost instantaneously, that in the floating-body device (dependent on the carrier generation lifetime) is very long compared to operating periods of high-fre-

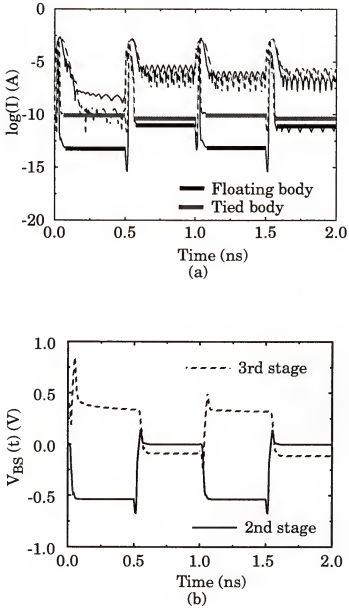


Fig. 5.12 SOISPICE predictions of the three-stage NFD/SOI CMOS inverter chain; L = 0.2 μm, V_{DD} = 1.5 V; (a) transient currents in the middle stage of the inverter chain with floating and tied bodies; (b) the floating-body voltage of the n-channel transistors in the chain. The physical off-state leakage (channel) currents are highlighted in (a); the noisy currents reflect errors associated with the integration method used in SPICE.

quency signals in scaled CMOS circuits; the depletion charge can only be neutralized by the reverse generation currents of the source/drain-body p-n junctions in the floating-body device. Thus, $V_{BS} < 0$ is predominant during the off-state period of the floating-body device, and the current is suppressed because of the implied higher threshold. The static power in the circuit is effectively decreased by this current "undershoot".

However, this "undershoot" effect shows a hysteretic behavior; the quantitative amounts of the undershoot are different in time as obvious in Fig. 5.12(a). The hysteresis is somewhat exaggerated in Fig. 5.12, due to the way the initial conditions are determined for the simulated circuit in SOISPICE; off- and on-transistors are assumed to have reached steady-state charge conditions under their respective initial steady-state bias conditions. In contemporary operation of fast circuits, continuous switching with reasonable duty cycle dictates that the device will operate with its floating-body charge depleted because the generation of the carriers (charging) is generally slower than the recombination (discharging). Therefore, the dynamic current undershoot will prevail in the circuit operations. This "undershoot", accompanied by the speed enhancement in Sec. 5.2.2, constitutes unique benefit of floating-body circuits, which will be more eminent in low-voltage/power applications.

An "overshoot" in contrast to "undershoot" occurs as the MOSFET is switched on from the off state. If the device reaches the steady-state in the off state, stored majority carriers become excess carriers upon the subsequent

switching and remain in the body until they recombine. The concomitant $V_{BS}(t) > 0$ lowers the threshold voltage, thereby producing a transient drain current overshoot [Lim84]. The magnitude of the overshoot depends on the switching frequency which will define how efficiently the majority carriers are replenished in the off state by carrier generation. In contemporary SOI technology, the devices are switched very fast, and typically the off-state time is much shorter than the generation lifetime. Hence the carriers are not significantly replenished, and the overshoot is usually negligible. SOISPICE simulations of CMOS inverter circuits, using representative carrier lifetimes to define I_R in (2.66) and I_{Gt} in (2.67), confirm this and predict no significant reduction in propagation delay due to this overshoot unless the supply voltage is very low and off-state times are relatively long as in the triple-pulse experiment described in Sec. 5.3.1.

5.4 Dynamic Instabilities due to Hysteresis in CMOS Circuits

The V_T hysteresis must be more carefully investigated because it can cause problems in some CMOS circuit applications, especially where some of the constituent floating-body NFD MOSFETs can randomly sit for a very long time, allowing the devices' body charge to be fully charged or discharged. The hysteretic V_T variation is maximized in such cases. Using the conventional design methodology based on the device/circuit DC characteristics without considering the V_T variation can result in circuit instabilities.

The effect of V_T hysteresis on circuit performance is implied in Fig. 5.13 where SOISPICE-simulated transient voltage-transfer characteristics of an NFD/SOI CMOS inverter are shown. For the characteristics in Fig. 5.13, the inverter input voltage is ramped up and down, rather than being applied steadily, at a very slow rate so as to obviate the gate propagation delay but fast enough to see the hysteretic effects defined by the recovery of the body charge. Both the forward- and reverse-swept transfer characteristics are offset from the steady-state one because of the body charging-induced V_T lowering of the transistor being switched from off to on. The ramp rate in the simulations is 1 V/μs, which suggests that again the floating-body charge recovery is not fast enough to settle itself to DC conditions. Therefore, the inverter characteristics in real high-frequency circuit operations will not be identical to the DC case.

Hence, the hysteresis casts doubts on the direct applicability of the established circuit design methodology to the NFD/SOI circuit design. In the conventional bulk CMOS circuit design, the hysteresis does not exist, and the circuits could be designed based on the representative DC characteristics of their constituent devices. A circuit designer, however, must recognize the hysteresis in designing NFD/SOI CMOS circuits, especially when the circuits operate in a random fashion, maximizing the threshold variation in the circuit. Indeed the hysteresis in the simulated characteristics of Fig. 5.13 is for the worst case. The devices are allowed to stay in one state infinitely before the ramping of the input voltage starts, and hence, the hysteresis is maxi-

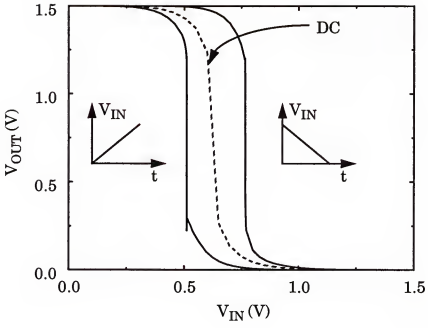


Fig. 5.13 SOISPICE-simulated forward- and reverse-swept (slow-) transient voltage-transfer characteristics of a floating-body NFD/SOI CMOS inverter. The steady-state characteristic is shown for comparison. The ramp rate is 1V/μs.

mized. Nevertheless, Fig. 5.13 suggests that a floating-body circuit will experience reduced noise margin relative to that expected based on DC characteristics, thus possibly resulting in instability under adverse circumstances.

Such a case could be encountered in an SRAM (static random access memory) cell in Fig. 5.14. To demonstrate the instabilities in the circuit operation, we simulated a "read-0" operation of a typical SRAM circuit. The bit-lines (BL and $\overline{BL}$) in the simulations are pre-charged to 2/3 of V_{DD} where V_{DD} is 1.5 V in this 0.2 μm circuit. As the cell is selected by raising the word line (WL) up, the bit line (BL) on the 0-stored side discharge through the n-channel transistor (M1), while the other side ($\overline{BL}$) reaches V_{DD} through the p-channel transistor complementarily. For a stable operation, SRAM should be designed such that when V_{WL} goes high, a momentary V_{GfS} on M2, defined by the on-resistance of M1 and M3 in Fig. 5.14, should not exceed V_T of M2 minus a pre-defined noise margin.

Figure 5.15 shows predicted waveforms on the bit line (V_{BL}) for a "read-0" operation in both floating- and tied-body SRAM circuits. First, under the normal word-line voltage ($V_{WL} = V_{DD} = 1.5$ V), both types behave normally, with the floating-body circuit faster than the other. This again reveals the floating-body benefit. But under noisy situations ($V_{WL} = 2.0$ V > V_{DD}), the floating-body circuit fails to operate, while the tied-body circuit remains stable and is actually faster than under normal V_{WL}. This instability is due to the reduced noise margin caused by the hysteretic threshold voltage reduction of the off-transistors (M2 and M3). As soon as V_{WL} goes high to select the cell, a

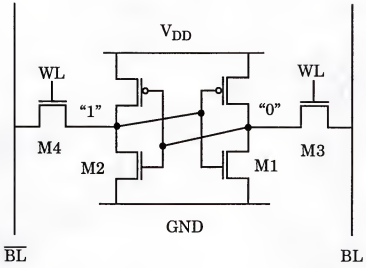
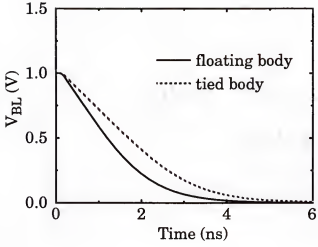
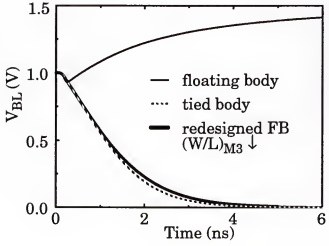


Fig. 5.14 Simulated SRAM circuit with 0 written and pre-charged bit lines (BL and $\overline{BL}$)



(a)



(b)

Fig. 5.15 SOISPICE-predicted bit-line (BL) waveforms during "READ-0" operation of the SRAM circuit; (a) under normal word-line voltage ($V_{WL} = V_{DD}$), (b) under noisy word-line voltage ($V_{WL} > V_{DD}$).

fraction of precharged bit-line voltage, defined by the resistive division of M1 and M3, is applied at the front gate of M2. That voltage tends to get higher in the floating-body circuit than the tied-body circuit because of the lowered threshold-voltage and the concomitant on-resistance reduction of M3. This, in combination with the threshold reduction in M2, which was off for the same period as M3, makes it even easier for M2 to be turned on during the read operation. To remedy this problem and stabilize the circuit, the floating-body circuit has to be redesigned to increase the on-resistance of M3 by reducing its aspect ratio (W/L), which obviously degrades the circuit performance as it can be evidenced in Fig. 5.15.

Another example of this type of possible instability can be found in a sense amplifier. Figure 5.16 shows the simulated circuit. The sense amplifier is an important circuit block in memory devices which detects a small differential signal taken on the complementary bit lines, and amplifies it on high capacitive bit lines to respective logic states in a sufficiently short time. Connected to DRAM (dynamic random access memory) cells as in Fig. 5.16, the sense amplifier also refreshes the selected memory cell at the same time. Figure 5.17 shows the SOISPICE-predicted bit-line waveforms in sense amplifier operations. If we concentrate first on the tied-body characteristics in Fig. 5.17(a), we see that the sense amplifier does sense and amplify in a stable manner cycle after cycle. However, in the floating-body circuit, it becomes unstable, and actually flips the data after a few cycles. We again attribute this to the hysteretic floating-body voltage instability, as evidenced in Fig. 5.17(b)

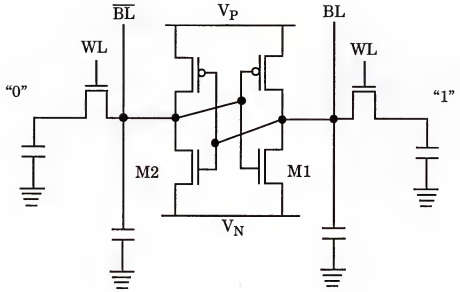
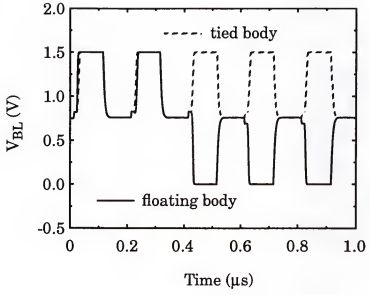
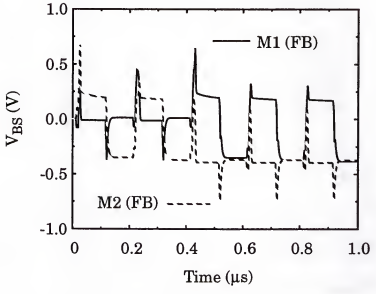


Fig. 5.16 Simulated sense amplifier circuit to exemplify floating-body instability. DRAM cells storing complementary data are attached at both ends.



(a)



(b)

Fig. 5.17 SOISPICE-simulated sense amplifier operation versus time; (a) bit-line voltage, (b) floating-body voltage of the coupled pair, M1 and M2.

where the floating-body voltages of M1 and M2 are plotted. M1 and M2 constitute a bistable coupled pair. Even though the signals applied to the gates of M1 and M2 are normal ($V_{GfS}(M2) > V_{GfS}(M1)$ in the simulations due to stored data), the hysteretic $V_{BS}(t)$ affects the actual gate drive ($V_{GfS} - V_T(t)$). In some cases, this hysteretic variation is not enough to cause an upset; the circuit performance is only affected as a matter of timing. The relative shift in position of the floating-body bit-line waveform for the first two cycles in contrast to the stable tied-body waveform in Fig. 5.17(a) shows this. However, the hysteretic difference can be large enough at one point to override a differential signal applied on the gates. This flips the data while amplifying as exemplified in Fig. 5.17. To avoid such instabilities in the sense amplifiers, body-ties are recommended [Hwa91]. However, they could be prohibitive in terms of technology complexity and cost.

5.5 Conclusions

In this chapter, we investigated the floating-body effects in the NFD/SOI CMOS device and circuits, using SOISPICE into which our new physical and compact model is implemented. The simulations provided the needed design insights regarding the floating-body effects. The subthreshold kink results due to the non-ideality of the junction, but the purported benefit is intangible and should be suppressed due to the implied high off-state current. Two floating-body effects (I_{Gt}-driven V_T reduction and off-state current suppression due to dynamic V_T increase) were shown to be beneficial to NFD/SOI

CMOS circuits via inverter-chain simulations. Supported by an experiment, the SOISPICE simulations showed that the floating-body charge dynamics govern the hysteretic V_T-fluctuation during circuit operation. This hysteresis can cause severe instabilities in floating-body NFD/SOI CMOS circuits, exemplified by SRAM and sense amplifier simulations using SOISPICE. The remedy to the instability due to hysteresis would seem to be design conservatism and/or technology complexity, which negate the beneficial effects.

CHAPTER SIX
SOI DRAM DESIGN AND ASSESSMENT

6.1 Introduction

Recently, the SOI technology has gained much interest for application to the next generation high-density DRAM, possibly replacing the established bulk-silicon technology. Technological merits from simpler processing aside, the device built on an SOI substrate implies electrical advantages in DRAM applications; soft-error immunity, smaller bit-line capacitances, and reduced leakage current due to smaller junction area all contribute toward a smaller cell capacitance and thus smaller cell size and higher density, purportedly without degradation of the data-retention characteristics of the cell [Tan94, Nis94]. A successful DRAM on SOI with as high a density as 16 megabits has been reported [Kim95].

However, there is a concern about dynamic data retention characteristics of DRAMs on SOI, although the underlying physics is not unequivocal [Mor95]. For SOI DRAM cells, the transistors are usually partially depleted (PD) for an acceptable threshold voltage in order to suppress the leakage current, and their bodies are left floating for high densities; the cell transistor is hence prone to unstable floating-body operation. For example, because of the bit-line voltage fluctuation due to the dynamic operations on adjacent cells,

the floating-body $V_{BS}(t)$ instability occurs, possibly resulting in an increase of I_{DS} of the cell transistor. Data retention time in dynamic situations can thus be affected in a peculiar way, unlike in bulk cells where steady-state leakage current defines retention. For a floating-body SOI DRAM to be viable in future giga-bit chips and beyond, these instabilities of data retention time must be avoided to guarantee stable operation.

In this chapter, we use our physical model for the NFD/SOI MOSFET in SOISPICE to characterize in detail the dynamic retention and to demonstrate how a periodic body-charge removal effected by the data refresh ensures extremely long data retention time suitable for 1-gigabit SOI DRAM.

6.2 Mechanisms for Possible Data Retention Degradation in DRAM Cell

Figure 6.1 shows the simple DRAM cell circuit, used in simulations throughout this chapter. The SOISPICE charge-based model (refer to Fig. 3.1 for its network representation) accounts for the capacitive couplings in the floating-body NFD/SOI MOSFET, governed by the nodal equation at the floating-body node. When $V_{GfS}(t)$ swing occurs, either by V_{BL} fluctuation or V_{WL} rise, dQ_B/dt is effected, depending on the $Q_B(t)$ and $V_{GfS}(t)$. The body charging current dQ_B/dt, with the generation and recombination currents in the device, defines $V_{BS}(t)$, which can be positive or negative. The $V_{BS}(t)$ defines the transient leakage current through the device, $I_{leak}(t)$. Data retention time of a DRAM cell is a function of this current and the amount of the charge stored in the storage capacitor, C_S. C_S cannot be increased arbitrarily due to area

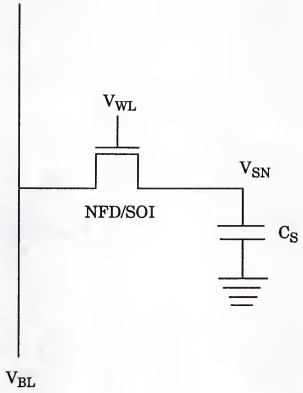


Fig.6.1 NFD/SOI DRAM circuit.

restrictions, and thus I_{leak} has to be controlled to achieve the best retention time for the cell.

If the cell transistor is a bulk device, the body terminal can be tied to a desired potential; then the leakage current through the device is constant in time depending on the threshold voltage of the device. However, due to the unique floating-body charge dynamics, the leakage current of a floating-body cell transistor is not constant, but varies dynamically. Figure 6.2 shows SOISPICE-predicted data decay of the DRAM cell in Fig. 6.1 with typical model parameters, for both tied- and floating-body MOSFETs. In the former case, the body is tied to ground, at which the bit line is also held in the simulation. Whereas the constant decay rate implies a constant I_{leak} in tied-body (bulk-like) circuit, the decay rate in the floating-body cell varies in time, according to the floating-body voltage $V_{BS}(t)$. Due to the depletion charge uncovered by recombination while the gate is high (for writing and/or refresh), $V_{BS}(t)$ becomes negative when the NFD/SOI device is turned off (a negative V_{GfS} swing). The consequential V_T increase suppresses the leakage current and thus the lower decay rate prevails initially in the floating-body cell. However, as time elapses, the depletion charge and negative $V_{BS}(t)$ diminish in magnitude as the carriers, created in the device body mainly by thermal generation mechanisms, neutralize the depletion charge. If the nominal (DC) V_T of the cell transistor is not designed sufficiently high, the cell can fail to retain the stored data because of possible V_T reduction from the nominal value due to a positive $V_{BS}(t)$. If $V_T(t)$ is reduced such that I_{leak} is weak-inversion chan-

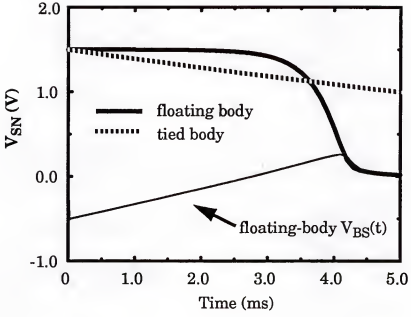
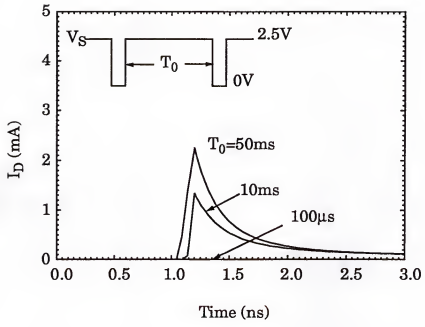


Fig.6.2 SOISPICE-predicted decay of the stored data voltage (V_{SN}) in floating- and tied-body PD/SOI DRAM cells.

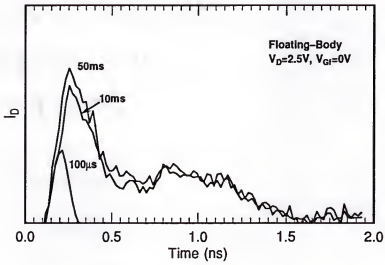
nel current instead of generation leakage current, the decay now becomes very abrupt, as indicated in Fig. 6.2, because I_{leak} increases exponentially with $V_{BS}(t)$ which is increasing in time (refer to subthreshold characteristics in Fig. 3.7.(b)).

Another failure mechanism is possible when V_{BL} varies fast. When V_{BL} is maintained above V_{WL} ($V_{GfS} < 0$), the device is driven to accumulation. The accumulation charge in the body has to be removed from the body when a positive V_{GfS} swing is effected by the fluctuation on the V_{BL} (i.e., V_{BL} going from high to low). Depending on the degree of accumulation and the rate of the swing, this fluctuation can instigate a significant transient parasitic-BJT current, $I_T(t)$ [Pel95], dQ_B/dt being the effective base current. Supported by the experiment in [Pel95], Fig. 6.3 shows the SOISPICE-predicted transient drain current under such a V_{BL} swing with the gate held at 0V; the current is predominantly BJT current. The magnitude of the induced BJT current depends on dQ_B/dt, which is governed by the circuit design ($V_{GfS}(t)$ and dV_{GfS}/dt) and the device properties (dQ_{Gf}/dt, dQ_S/dt, I_R, and I_{Gt}). Since the time required for the dynamic body discharging can be relatively long ($<Q_B/I_R> \sim$ 1ns), the charge transported by the $I_T(t)$ pulse can be substantive (~ 100fC/ μm) and could be enough to discharge C_S (~ 20fF).

These effects, however, are also hysteretic. As evident in Figs. 6.2 and 6.3 and analogous to the results of the triple-pulse gate delay experiment in Sec. 5.3.1, the severity of the effects depends on the off period of the operation during which $V_{BS}(t) < 0$ and the free carriers are supplied by thermal



(a)



(b)

Fig.6.3 Transient drain currents of an NFD/SOI MOSFETs, following the 2.5-to-0V source transition with V_D = 2.5V for varying time (T_0) between pulses; (a) SOISPICE-predicted current; (b) measured current. t_{ox} = 7nm, t_{Si} = 140nm, L_{eff} = 0.3µm, W = 10µm.

generation, either to neutralize the depletion charge or to accumulate the surface. For a viable floating-body NFD/SOI DRAM cell, such faulty occurrences described above have to be avoided by not allowing enough time for recovery of substantive Q_B, thereby keeping $V_{BS}(t)$ low. Fortunately, the time required to charge the body is extremely long in a typical contemporary device ($<Q_B/I_{Gt}>$ ~ 10ms), and hence it is conceivable to design SOI DRAMs with periodic body discharging effected by the data refresh to maintain $V_{BS}(t) < 0$ with $V_{Gf} = 0$.

6.3 Dynamic DRAM Operation

We used SOISPICE to simulate a representative low-voltage dynamic operation of the DRAM cell. The NFD/SOI MOSFET assumed has W/L = 0.2μm/0.2μm and t_{ox} = 7nm; its simulated DC subthreshold characteristics are plotted in Fig. 6.4 for low and high (V_{DD} = 1.5V) V_{DS}. The characteristics in Fig. 6.4 do not alone indicate the data retention capability. The V_{GfS} = 0V leakage current is high (I_{CH}), but the effective off-state current in DRAM cell comprises the dynamic $I_{CH}(t)$ and $I_T(t)$. The result of the DRAM simulation with C_S = 20fF is plotted in Fig. 6.5. The key to understanding the peculiarities of the dynamic retention characteristics is $V_{BS}(t)$, which as discussed above depends on $Q_B(t)$, I_R, and I_{Gt}. At t = 5ns, V_{BL} drops to 0V from 0.75V (=V_{DD}/2), defining a positive $V_{GfS}(t)$ pulse. Thus, $V_{BS}(t)$ increases and then decays as the excess holes recombine in accord with $I_R(V_{BS})$. During this period, discharging of C_S is accelerated as reflected by the predicted $V_{SN}(t)$. This discharge is due mainly to the induced $I_T(t)$, although generally the

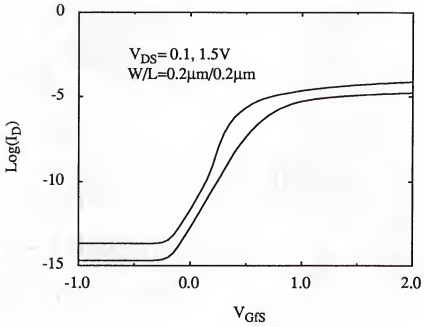


Fig.6.4 SOISPICE-simulated subthreshold characteristics of NFD/SOI nMOSFET used in the floating-body DRAM cell simulations of Fig. 6.5.

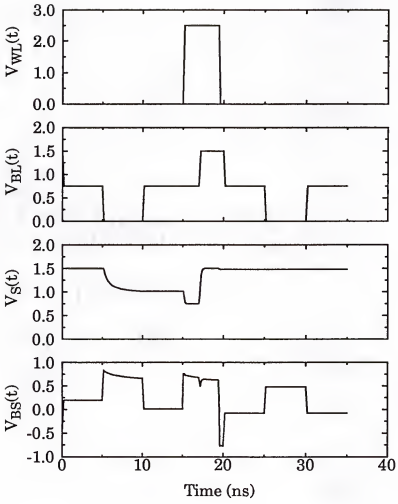


Fig.6.5 SOISPICE-simulated dynamic operation of floating-body NFD/SOI DRAM cell

implied negative shift of $V_T(t)$ could possible mean significant $I_{CH}(t)$ as well. These bit-line transients underlie the reduced dynamic (vs. static) data retention time as seen in [Mor95]. At t = 10ns, V_{BL} is pulsed up to $V_{DD}/2$, thereby dropping $V_{BS}(t)$ to a lower level than at t = 5ns, because of hole deficiency due to the body discharge which occured concomitantly with $I_T(t)$. At t = 15ns, V_{WL} is boosted up to 2.5V, and subsequently V_{BL} is pulsed to V_{DD}, effecting a refresh (or a write-1) operation as reflected by the predicted $V_{SN}(t)$. Again, $V_{BS}(t)$ first increases significantly because of the excess holes, and then falls off since I_R is high. The effect of (Q_B) discharge is evident at t = 20ns when V_{WL} drops back to 0V; $V_{BS}(t)$ goes negative due to the hole depletion which is not supported by $V_{GfS}(t)$. When V_{BL} subsequently drops back to $V_{DD}/2$, $V_{BS}(t)$ increases but remains lower than its value prior to the V_{WL} pulse, reflecting an effective discharge. Consequently when V_{BL} is pulsed down and then up again, the resulting positive peak in $V_{BS}(t)$ is smaller than that at t = 5ns, and hence the discharging due to $I_T(t)$ is substantively reduced. The complicated nature of the transistor's I_{leak} and the dynamic data retention are thus evident. To ensure long enough retention time, the body charge discharge effected during write/refresh must be sufficient to keep $V_{BS}(t)$ low enough. Note that underlying this hysteretic $V_{BS}(t)$ are $Q_B(t)$ and $V_{GfS}(t)$. Therefore, the best way to overcome this complexity is to focus on the floating-body charge, and its variation.

6.4 Floating-Body Charge Variation under Refresh

The dynamic body charging current, dQ_B/dt, in the floating-body NFD/SOI MOSFET must be balanced by the recombination and the generation of the device to satisfy the nodal equation at the floating-body node in Fig. 3.1. However, depending on the floating-body charge condition of the device, either recombination or generation is dominant at a point in time. For example, when an excess body-charge condition occurs in the device due to a positive V_{GfS} pulse, the recombination is predominant while $V_{BS}(t) > 0$, removing the excess carriers in the body. For a negative pulse, charging due to generation is predominant, while $V_{BS}(t) < 0$. Therefore, reasonable approximations of the complex body-charge dynamics are

$$\frac{dQ_B}{dt} \cong -I_R \qquad \text{when } V_{BS}(t) > 0, \tag{6.1}$$

$$\frac{dQ_B}{dt} \cong I_{Gt}. \qquad \text{when } V_{BS}(t) < 0. \tag{6.2}$$

Characterizing (6.2) is simple because I_{Gt} is relatively independent of biases. Characterizing (6.1), however, involves a very complex nature of the device due to the various components of I_R and Q_B, which dictates different time constants [Lim84]. As an illustration, Fig. 6.6 shows schematic characteristics of Q_B discharging and charging characteristics in a DRAM cell, which are approximately governed by (6.1) and (6.2) respectively. The Q_B-discharging characteristic is exponentially fast when $V_{BS} > 0$ is large (BJT-charge domi-

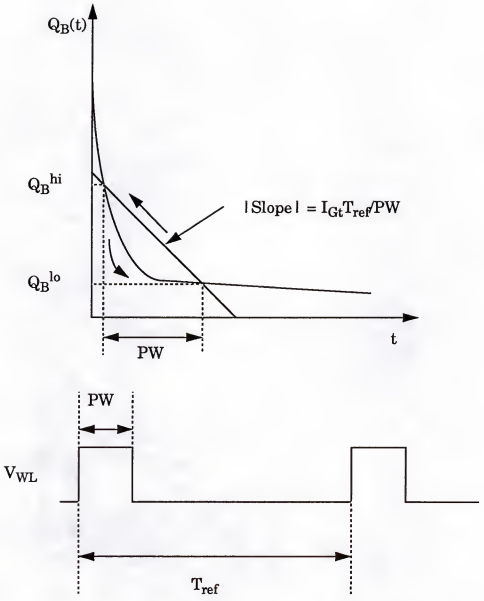


Fig.6.6 Schematic illustration of floating-body charging/discharging characteristics versus time under repetitive V_{WL} pulse.

nant), but is slowed later in time (MOS-charge dominant) with reduced $V_{BS} > 0$. The charging characteristic is slow but linear, as implied by (6.2). These principles regarding the floating-body charge dynamics apply to any floating-body circuit. In dynamic DRAM operation, however, the characterization can be facilitated by the well defined word-line voltage, V_{WL}. For refreshes of the stored data, the V_{WL} pulse is applied to the gate of the cell transistor periodically, and the counteracting charge dynamics of (6.1) and (6.2) alternate regularly as dictated by V_{WL}. The regularity can be disrupted by a possible discharge due to the bit-line fluctuation while $V_{WL} = 0$, as shown in Fig. 6.5, and to a V_{WL} rise between refresh for a read/write operation on the cell. However, we assume the regularity in the subsequent discussion, which allows us to consider the worst-case floating-body charge condition for DRAM operations.

To present both charging and discharging characteristics on the same plot, the charging slope is scaled depending on the duty cycle of V_{WL} and the refresh time, T_{ref}, as indicated in Fig. 6.6. The scaling effectively accounts for the charging that occurs for the longer period of time when $V_{GfS} = 0$. When the V_{WL} pulse turns the device on, $Q_B(t)$ discharges to a certain value from which the charging starts while $V_{WL} = 0$. The $Q_B(t)$ to which the device charges during $V_{WL} = 0$ is the value from which it discharges at the next V_{WL} pulse. If $Q_B(t_n)$ at $t = t_n$ (time when n-th V_{WL} pulse is applied) is higher than Q_B^{hi} in Fig. 6.6, the discharge will be faster than from Q_B^{hi}, and therefore $Q_B(t_{n+1})$ will be smaller than $Q_B(t_n)$ for a given I_{Gt} and V_{WL}. The opposite

applies when $Q_B(t_n)$ is lower than Q_B^{hi}. Therefore the general transient will tend to settle at Q_B(hi) and Q_B(lo) in Fig. 6.6. Then we can define regular-state $Q_B(t)$'s before and after the V_{WL} pulse as the intersections of the two characteristics, i.e., Q_B^{hi} and Q_B^{lo}.

The higher Q_B^{hi} is, the higher $V_{BS}(t)$ will be for a given positive V_{GfS} pulse, effected either by V_{WL} or by V_{BL}. If $V_{BS}(t)$, defined by Q_B^{hi} in Fig. 6.6 and subsequent fluctuation of V_{BL}, does not cause the unstable increase of I_{leak} as described in Sec. 6.2, and if the refresh time, which defines with I_{Gt} the charging slope, is short enough, then DRAM operation will be stable. Note that longer T_{ref} implies steeper charging slope; thus it is probable that Q_B^{hi} gets higher. Hence, it is suggested that the maximum allowed refresh time depends on the generation capability in the device.

Note that the discussions regarding the charge dynamics in Fig. 6.6 consider the worst case in DRAM operations. If V_{WL} is pulsed between refreshes for write/read operation, $Q_B(t_n)$ will be lower than Q_B^{hi} due to the discharging effected by V_{WL}. V_{BL} fluctuation, even without discharge induced by it, also means that $Q_B(t_n)$ will be lower than Q_B^{hi}; the charging does not continue after an equilibrium dictated by lower V_{BL} is reached. Consequently the $V_{BS}(t)$ that will develop transiently according to $Q_B(t)$ will be smaller in both these cases than the illustrated worst-case situations.

The goal for an optimal SOI DRAM is then obvious: it is to maximize the charging time by reducing I_{Gt} minimally, while designing a device to retain the stored data as long as possible by keeping $V_{BS}(t)$, defined by Q_B(hi)

and V_{BL} pulses, low. Then an indication of the required refresh time can be given as

$$\Delta t \cong \frac{\Delta Q_B}{I_{Gt}}, \tag{6.3}$$

which is defined by the carrier thermal generation lifetime, τ_G, as implied by (6.2) and Fig. 6.6.

6.5 DRAM Cell Transistor Design

The projected required refresh time for a 1-Gb DRAM is 1 second [Ito90]. Hence Δt given by (6.3) should be greater than 1 sec. ΔQ_B is dependent on various device parameters (e.g., t_{ox}, I_{R0}) and circuit variables (V_{DD}), but can be roughly estimated as

$$\Delta Q_B \sim C_{ox} W L V_{DD}. \tag{6.4}$$

Equation (6.4) yields for a typical 0.2μm technology about 0.1fC, which is compatible with SOISPICE predictions. Figure 6.7 shows the simulated floating-body discharge characteristics of a DRAM cell (0.2μm/0.2μm) with V_{WL} turned on ($V_{WL} > V_T+V_{DD}$) for varying model parameters. In all parameter variations, ΔQ_B is roughly identical to the estimation in (6.4); the discharge rate and $Q_B{}^{lo}$ are only affected by small factors. Then the required τ_G, from (6.3), (6.4), and (2.67), is estimated to be 0.1 μs, which is not incompatible with the evolution of SOI materials.

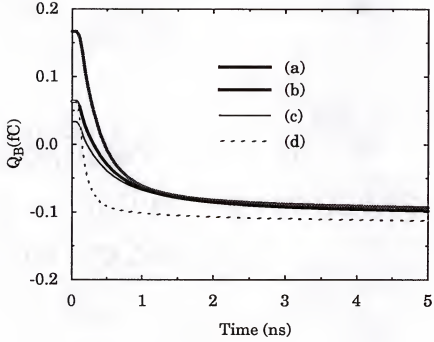


Fig.6.7 SOISPICE-simulated floating-body Q_B-discharge characteristics; (a) nominal device in Fig. 6.4 (t_{ox} = 7nm, JRO = 3E-9, WKF = -1.023 (n^+-poly)); (b) WKF=-0.5; (c) t_{ox} = 10nm; (d) JRO = 3E-8.

The needed generation time being defined, the NFD/SOI DRAM transistor should be designed to withstand $V_{BS}(t) > 0$ or to keep $V_{BS}(t)$ low enough. Therefore the nominal V_T of the device should be high enough to maintain $I_{leak} = I_{Gt}$ in spite of the V_T reduction effected by positive $V_{BS}(t)$ for most of the operation time. The most vulnerable case is when $V_{BL} = 0$. When $V_{BL} = V_{DD}/2$, the effective $V_{GfS} < 0$ can counteract a positive $V_{BS}(t)$. Note that $V_{BS}(t) > 0$ is due to the excess charge unaccountable by terminal voltage, V_{GfS}. The higher V_T implies the higher $Q_B(DC)$ for $V_{GfS} = 0$ V. Therefore, for a $Q_B{}^{hi}$ determined by the V_{WL} pulse and I_{Gt}, $V_{BS}(t)$ will be smaller in a high-V_T cell transistor due to smaller $Q_B{}^{hi} - Q_B(DC)$. For the same reason, the severity of the parasitic-BJT will dwindle. According to the characteristics in Fig. 6.4 and the $V_{BS}(t)$ in Fig. 6.5, the needed V_T increase would seem to be about 0.5 V. Increasing V_T that much by increasing doping will fail to achieve the desired shift of the subthreshold characteristics; the increased depletion body capacitance due to higher doping tends to degrade the subthreshold slope. Therefore a different gate material than n^+-poly, which was used in Figs 6.4 and 6.5, seems to be necessary to give the desired V_T shift.

Figure 6.8 shows the $V_{BS}(t)$ and $Q_B(t)$ from the simulated DRAM operation in Fig. 6.5, but includes higher-V_T results as well, effected by setting the front work function parameter, WKF, equal to -0.5 V (a la mid-bandgap gate). The same $Q_B(0)$'s in both simulations are enforced by initial $V_{BS}(0)$'s, which are lower in the higher V_T device. In the simulation result of Fig 6.8, $V_{BS}(t)$ in the high-V_T device is maintained lower as discussed above,

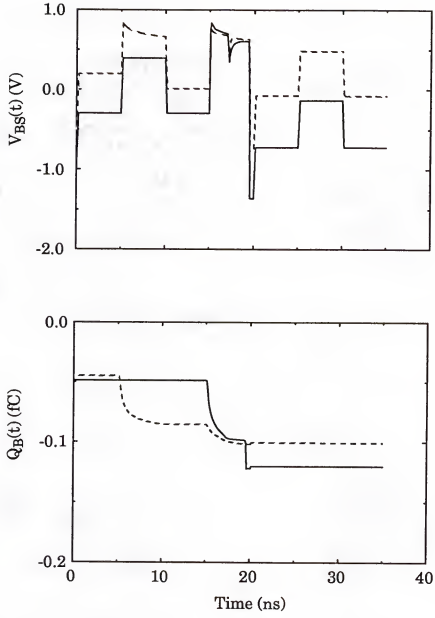


Fig.6.8 SOISPICE-predicted dynamic floating-body voltage and charge during dynamic DRAM operations in Fig. 6.5. High-V_T device (solid lines) and the device of Fig. 6.4 (broken lines) are compared.

inhibiting $I_T(t)$ and $I_{CH}(t)$ for the same V_{BL} swing. Therefore the high-V_T cell will retain V_{SN} as well as Q_B in a more stable way until the next refresh is effected. The retention of Q_B in the high-V_T cell is apparent at t = 15 ns when V_{WL} turns on the device. During the V_{WL} pulse, higher $V_{BS}(t)$ and consequently faster discharge are induced due to the higher $Q_B(t_n)$ in the high-V_T cell. Therefore, the Q_B's of both cell transistors do not differ significantly regardless of V_T, and the ΔQ_B's over the simulation period are comparable.

Now the DRAM operation with the high-V_T cell will be stable; no $I_T(t)$ and $I_{CH}(t)$ are effected, due to the deficiency of holes. Until the next refresh, $Q_B(t)$ as well as $V_{BS}(t)$ will increase as the generated carriers again charge the body. As a crosscheck, the time to reach $V_{BS}(0)$ in Fig. 6.8 again can be estimated from (2.48), (2.75), (2.67), and ΔV_{BS} in Fig. 6.8 as

$$\Delta t \cong \frac{WLC_{ox}\Delta V_{BS}}{I_{Gt}} \cong 2 \text{ sec}, \tag{6.5}$$

with τ_G taken from the estimation in (6.4). Therefore we can expect a stable data retention using this high-V_T cell.

To effect a full-V_{DD} write in DRAM, V_{WL} must be greater than $V_{DD}+V_T(on)$. Therefore the high threshold voltage portends a power consumption problem in DRAM; the needed boosted V_{WL} means greater power dissipation on the word-line, which is proportional to $C_{WL}V_{WL}^2$. However, in the NFD/SOI DRAM, a V_T reduction results during write/refresh, which alleviates the requirement of a boosted V_{WL} level due to the excess charge as implied by $V_{BS}(t) > 0$ in Figs. 6.5 and 6.8 during write/refresh. The recombina-

tion of the excess carriers are not fast enough to fully discharge the body during the write/refresh period, and therefore the Q_B^{lo} may not necessarily be the on-state Q_B(DC).

Figure 6.9 shows the effect of V_T reduction where $V_{SN}(t)$ is plotted for tied- and floating-body cells, with $V_{BL} = V_{DD}$ to effect a write-1. The significant difference of V_{SN} between floating- and tied-body stresses the effect of threshold voltage. Due to the above mentioned V_T reduction, the V_{WL} in the NFD/SOI DRAM need not be above V_{DD} by the nominal V_T, whereas that in the bulk device has to be boosted even higher due to the body effect by $V_{BS} < 0$. The $V_{BS} > 0$ in Fig. 6.8 can be translated into about 0.2 V of V_T reduction, which is not very sensitive to the device variations according to Fig. 6.7. Hence SOI DRAM is even more attractive for low-power applications.

6.6 Conclusions

Using SOISPICE simulations, the mechanisms that can result in abnormal dynamic retention time in floating-body NFD/SOI DRAM cell were investigated. The short dynamic retention time is due to $I_T(t)$ and $I_{CH}(t)$ induced by positive $V_{BS}(t)$, which is defined by Q_B, and V_{GfS}. Due to the regular data refresh, essential to the DRAM operation, the body charge, Q_B, is periodically discharged, which could suppress these occurrences. Therefore, the DRAM design to control V_{BS} by periodic discharge was assessed. The floating-body charge condition depends on the generation rate in the device and the refresh time of the DRAM. The difference between $Q_B(t)$ and the steady-

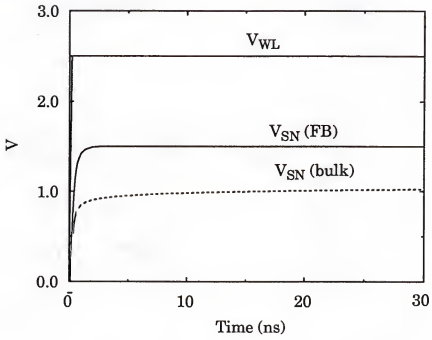


Fig.6.9 SOISPICE-simulated write-1 characteristics of high-V_T cell transistor for floating- and tied- (bulk like) body connections.

state body charge, $Q_B(DC)$, defines $V_{BS}(t)$, and thus a high V_T (a la mid-band-gap gate) device results in lower V_{BS}, which is adequate for a 1-Gb DRAM given $\tau_G > 0.1$ μs. The resultant high V_T did not cause a significant increase in word-line power consumption due to the dynamic V_T reduction during write/refresh, thereby showing the viability of the NFD/SOI technology in the 1-Gb DRAM application.

CHAPTER SEVEN
SUMMARY AND SUGGESTIONS FOR FUTURE WORKS

7.1 Summary

This dissertation addressed the modeling of NFD/SOI MOSFETs and applications to high-performance/low-power ULSI design. The significant achievements made through this work can be summarized as follows.

First, with an objective to develop a useful tool that reflects the technology changes in the NFD/SOI technology, we extended and modified existing models to create a new charge-based physical compact NFD/SOI device model. The model includes the various short-channel effects, pertinent in the scaled devices. The regional approach, which characterizes the regions separately and interlinks them later by numerical methods, facilitated the development of the complete charge-based model.

Second, the model was implemented in SOISPICE as a new entry, which enabled its applications in numerous device/circuit analyses. To provide a solid foundation for further use of our model, the SOISPICE-predictions were verified against device simulation results and measurement data from differently scaled technologies. The model parameter-extraction schemes, based on the physical nature of our model, were used in the comparisons. They do not need a rigorous optimization, which separates our model from the

existing empirical ones. The predictions were excellent, reliably predicting the floating-body effects in NFD/SOI devices.

Third, the efficacy of body-ties were analyzed using our model. Aided by the SOISPICE simulations, simple and insightful breakdown conditions were derived, which later were discussed in connection with the device parameters, especially the parasitic body resistance. Body ties in the NFD device have to be carefully designed to obtain desired improvements (i.e., suppression of the floating-body effects) because of the high intrinsic body resistance due to the thin Si layer. The implied area penalty for an effective body tie in the NFD/SOI device could overwhelm the benefit.

Fourth, the floating-body effects in the NFD/SOI device and circuits were comprehensively characterized. SOISPICE device/circuit simulations provided the needed insights regarding the effects and their influence on NFD/SOI circuits for their pragmatic exploitation and/or suppression. The floating-body effects are complicated in transient circuit operations due to the floating-body charge dynamics that cause hysteretic threshold voltage variations as well as the transient BJT current. The V_T variation was supported by measured variations in the gate delay of inverter circuits, qualitatively predicted by SOISPICE simulations.

The kink effect, seemingly beneficial in DC I-V characteristics, was revealed to be not beneficial to circuit performance from inverter-chain simulation results and, in fact, can be detrimental because of implied high off-state current. Two floating-body effects (I_{Gt}-driven V_T reduction and off-state cur-

rent suppression due to dynamic V_T increase) were shown to be beneficial. However, hysteresis can cause severe instabilities in floating-body NFD/SOI CMOS circuits due to the floating-body charge dynamics that govern V_T fluctuation during circuit operations. The floating-body instabilities were exemplified by SRAM and sense amplifier simulations using SOISPICE. Remedy of the instability would seem to be design conservatism and technology complexity, which will compromise the beneficial effects.

Finally, with insights regarding the dynamic floating-body effects, the issues in NFD/SOI DRAM design were addressed. Due to the body charge fluctuation that yields $V_{BS}(t) > 0$, a significant and unstable increase in the leakage current of the cell transistor can occur during data retention. The suppression of these events could be achieved by frequent body-charge discharge, effected by regular data refresh. The key to a successful DRAM design was long generation lifetime and high V_T. The generation rate governed by the lifetime controls the needed frequency of the body charge refresh, which is in direct relation with the refresh time. High V_T keeps $V_{BS}(t)$ low by the implied high body-charge conditions in the DC off state. The hysteretic V_T reduction during refresh/write ameliorates the high-power consumption in DRAM, implied by the high V_T design. The needed generation lifetime in SOI materials is not incompatible with the emerging SOI technology and the projected refresh time for 1-Gb DRAM and beyond.

7.2 Suggestions for Future Works

First, the parameter-extraction procedure needs to be developed in a more systematic manner. Based on the scheme in Chapter 3, the new improved procedure must carefully investigate the correlations among model parameters in their extraction. The importance of dynamic floating-body effects also requires an optimization of several parameters based on transient experiment such as in Fig. 5.9.

Second, our parasitic-BJT model is insufficient to give insight in device scaling. The existence of the depletion region in the low-doped body portends a 2-D carrier transport problem, which may not be easy to characterize. However, an accurate model is needed, including 2-dimensional carrier transport. The core of the improved model can be the characterization of the carrier charges in the depleted body, which should be consistent with the improved BJT charge modeling. The importance of parasitic-BJT related characterization has been exemplified, especially in Chapter 6 where the parasitic-BJT is shown to be induced by transient excitation.

Thirdly, the complex nature of the body charge dynamics requires refinements in the generation and recombination current models. Although they are seemingly secondary in device's terminal characteristics, their control of the modulation of the floating-body charge in time is of great importance in circuit predictions, due to the sensitivity of V_T and I_T on V_{BS} which they define, as exemplified in Chapters 5 and 6.

Finally, including the temperature dependence in the model parameters will greatly improve the integrity of the model for CAD applications. The devices in a ULSI integrated circuit usually operate in elevated-temperature ambients due to the power dissipation by itself and in other parts of the circuit. The elevated temperature, for example, can accelerate the generation and recombination rates in the body, which are surely instrumental in defining the floating-body device characteristics. Therefore, simulations at representative ambient temperatures would be more meaningful. Such temperature-dependent modeling could further be the basis for adding an accounting for self-heating to SOISPICE

REFERENCES

[Bre80] J. R. Brews, W. Fichtner, E. H. Nicollian, and S. M. Sze, "Generalized guide for MOSFET miniaturization," IEEE Electron Device Lett., vol. EDL-1, pp. 2-4, Jan. 1980

[Bur89] R. L. Burden, and J. D. Faires, Numerical Analysis, 4th Ed., Boston: PWS-Kent, 1989

[Cho91a] J.-Y. Choi, Modeling and simulation of the fully depleted silicon-on-insulator MOSFET for submicron CMOS IC design, Ph.D. dissertation, University of Florida, Gainesville, FL, 1991

[Cho91b] J.-Y. Choi, and J. G. Fossum, "Analysis and control of the floating-body bipolar effects in fully depleted submicrometer SOI MOSFET's," IEEE Trans. Electron Devices., vol. 38, pp. 1384-1391, Jun. 1991

[Cod85] C. F. Codella, and S. Ogura, "Halo doping effects in submicron DI-LDD design," Tech. Digest International Electron Device Meet., pp. 230-233, Washington DC, Dec. 1985

[Col88] J. -P. Colinge, "Reduction of kink effect in thin-film SOI MOSFET's," IEEE Trans. Electron Devices, vol. 9, pp. 97-99, Feb. 1988

[Dav86] J. R. Davis, A. E. Glaccum, K. Reeson, and P. L. F. Hemment, "Improved subthreshold characteristics of n-channel SOI transistors," IEEE Electron Device Lett., vol. EDL-7, pp. 570-572, Oct. 1986

[Fit91] D. Fitzpatrick, SOISPICE User Guide, University of Florida, Gainesville, Florida, 1991

[Fos81] J. G. Fossum, and M. A. Shibib, "An analytical model for minority-carrier transport in heavily doped regions of silicon devices," IEEE Trans. Electron Devices, vol. ED-28, pp. 1018-1025, Sep. 1981

[Fos87] J. G. Fossum, R. Sundaresan, and M. Matloubian, "Anomalous subthreshold current-voltage characteristics of n-channel SOI MOSFET's," IEEE Electron Device Lett., vol. 8, pp. 544-546, Nov. 1987

[Fos93a] J. G. Fossum, P. -C. Yeh, and J. -Y. Choi, "Computer-aided performance assessment of fully depleted SOI CMOS VLSI circuits," IEEE Trans. Electron Devices, vol. 40, pp. 598-604, Mar. 1993

[Fos93b] J. G. Fossum, and S. Krishnan, "Current-drive enhancement limited by carrier velocity saturation in deep-submicrometer fully depleted SOI MOSFET's," IEEE Trans. Electron Devices, vol. 40, pp. 457-459, Feb. 1993

[Gar87] S. L. Garverick, and C. G. Sodini, "A simple model for scaled MOS transistors that includes field-dependent mobility," IEEE J. Solid-State Circuits, vol. SC-22, pp. 111-114, Feb. 1987

[Gre93a] K. R. Green, A model of the short-channel, metal-oxide-semiconductor, field-effect transistor for pragmatic mixed-mode device/circuit simulations, Ph.D. dissertation, University of Florida, Gainesville, FL, 1993

[Gre94b] K. R. Green, and J. G. Fossum, "A simple two-dimensional model for subthreshold channel-length modulation in short-channel MOSFET's," IEEE Trans. Electron Devices, vol. 40, pp. 1560-1562, Aug. 1993

[Hor92] A. Hori, A. Hiroki, M. Segawa, T. Hori, A. Shinohara, M. Yasuhira, and S. Akiyama, "High carrier velocity and reliability of Quarter-Micron SPI (Self-aligned Pocket Implantation) MOSFETs," Tech. Digest International Electron Device Meet., pp. 699-702, San Francisco, CA, Dec. 1992

[Hou95] T. Houston, private communications

[Hsu82] F. C. Hsu, P. K. Ko, S. Tam, C. Hu, and R. S. Muller, "An analytical breakdown model for short-channel MOSFET's," IEEE Trans. Electron Devices, vol. ED-29, pp. 1735-1740, Nov. 1982

[Hu93] C. Hu, "Silicon-on-insulator for high speed ULSI," Ext. Abstract of International Conference on Solid State Devices and Materials, pp. 137-139, Chiba, Japan, Aug. 1993

[Hwa91] J.-M Hwang, H. Lu, Y. D. Sheu, W. Bailey, P. Mei, and G. Pollack, "Premature breakdown in non-fully depelted SOI/MOSFET's with body-tied-to-source structure," Proc. International SOI Conference, pp. 34 -35, Vail Valley, CO, Oct. 1991

[Hwa93] J.-M. Hwang, E. Yee, T. Houston, and G. P. Pollack, "High performance submicron SOI/CMOS with an elevated source/drain structure," Proc. International SOI Conference, pp.132-133, Palm Springs, CA, Oct. 1993

[Ito90] K. Itoh, "Trends in megabit DRAM circuit design," IEEE J. Solid-State Circuits, vol. 25, pp. 778-789, 1990

[Jen94] K. A. Jenkins, J. Y.-C. Sun, and J.-L. Pelloie, "Measurement of SOI MOSFET I-V characteristics without self-heating," Proc. International SOI Conference, late news, Nantucket, MA, Oct. 1994

[Kim95] H.-S. Kim, S.-B. Lee, D.-U. Choi, J.-H. Shim, K.-C. Lee, K.-P. Lee, K.-N. Kim, and J.-W. Park, "A high performance 16m DRAM on a thin film SOI," 1995 Symposium on VLSI Technology of Tech. Papers, Kyoto, Japan, Jun. 1995

[Kri94] S. Krishnan, unpublished works on the nonlocal impact ionization

[Lim83] H.-K. Lim, and J. G. Fossum, "Threshold voltage of thin-film silicon-on-insulator (SOI) MOSFET's," IEEE Trans. Electron Devices, vol. ED-30, pp. 1244-1251, Oct. 1983

[Lim84] H.-K. Lim and J. G. Fossum, "Transient drain current propagation delay in SOI CMOS," IEEE Trans. Electron Devices, vol. ED-31, pp. 1251-1258, Sep. 1984

[Med92] MEDICI User's Manual, Technology Modeling Associates Inc., Palo Alto, CA, 1992

[Mor83] J. Ll. Morris, Computational methods in elementary numerical analysis, New York: Wiley, 1983

[Mor95] F. Morishita, K, Suma, M. Hirose, T. Tsuruda, Y. Yamaguchi, T. Eimori, T. Oashi, K. Arimoto, Y. Inoue, and T. Nishmura, "Leakage mechanism due to floating body and countermeasure on dynamic retention mode of SOI-DRAM," 1995 Symposium on VLSI Technology of Tech. Papers, Kyoto, Japan, Jun. 1995

[Mul77] R. S. Muller, and T. I. Kamins, Device Electronics for Integrated Circuits, New York; Wiley, 1977

[Nag92] R. Nagai, K. Umeda, and E. Takeda, "Low voltage, high gain, 0.2μm NMOSFETs by channel counter doping with As," Ext. Abstract of International Conference on Solid State Devices and Materials, pp. 695-696, Tsukuba, Japan, 1992

[Nis94] T. Nishihara, H. Moriya, N. Ikeda, H. Aozasa, and Y. Miyazawa, "Data retention in ultra-thin-film-SOI DRAM with buried capacitor cell," 1994 Symposium on VLSI Technology of Tech. Papers, pp. 39-40, Honolulu, HI, Jun. 1994

[Now93] E. J. Nowak, "Ultimate CMOS ULSI Performance," Tech. Digest International Electron Device Meet., pp. 115-118, Washington DC, Dec. 1993

[Pel95] M. M. Pelella, J. G. Fossum, D. Suh, S. Krishnan, and K. A. Jenkins, "Low-voltage transient bipolar effect induced by dynamic floating-body charging in PD/SOI MOSFET's," Proc. International SOI Conference, Tucson, AZ, Oct. 1995

[PIS89] PISCES User's Manual, Technology Modeling Associates Inc., Palo Alto, CA, 1989

[Pol91] G. Pollack, private communications

[Sha92] G. G. Shahidi, "SOI: new opportunities for sub-0.25μm VLSI," Ext. Abstract of International Conference on Solid State Devices and Materials, pp. 493-495, Tsukuba, Japan, Aug. 1992

[Sha93] G. G. Shahidi, T. H. Ning, T. I. Chappell, J. H. Comfort, B. A. Chappell, R. Franch, C. J. Anderson, P. W. Cook, S. E. Schuster, M. G. Rosenfield, M. R. Polcari, R. H. Dennard, and B. Davari, "SOI for a 1-volt CMOS technology and application to a 512Kb SRAM with 3.5ns access time," Tech. Digest International Electron Device Meet., pp. 813-816, Washington DC, Dec. 1993

[Su94] L. T. Su, J. E. Chung, D. A. Antoniadis, K. E. Goodson, and M. I. Flik, "Measurement and modeling of self-heating in SOI MOSFET's," IEEE Trans. Electron Devices., vol. 41, pp. 69-75, Jan. 1994

[[Sze81] S. M. Sze, Physics of Semiconductor Devices, 2nd Ed., Wiley, New York, 1981

[Tan94] T. Tanigawa, Y. Yoshino, H. Koga, and S. Ohya, "Enhancement of data retention time for giga-bit DRAMs using SIMOX technology," 1994 Symposium on VLSI Technology of Tech. Papers, pp. 37-38, Honolulu, HI, Jun. 1994

[Tsi82] Y. Tsividis, "Moderate inversion in MOS devices," Solid-State Electronics, vol. 25, pp. 1099-1104, 1982

[Vee88] S. Veeraraghavan, and J. G. Fossum, "A physical short-channel model for the thin-film SOI MOSFET applicable to device and circuit CAD," IEEE Trans. Electron Devices, vol. 35, pp. 1866-1875, Nov. 1988

[Yan92] R. -H. Yan, A. Ourmazd, and K. F. Lee, "Scaling the Si MOSFET: from bulk to SOI to bulk," IEEE Trans. Electron Devices., vol. 39, pp. 1704-1710, Jul. 1992

[Yeh93] P. -C. Yeh, and J. G. Fossum, "Subthreshold MOSFET conduction model and optimal scaling for deep-submicron fully depleted SOI CMOS," Proc. International SOI Conference, pp. 142-143, Palm Springs, CA, Oct. 1993

[Yeh94] P. -C. Yeh, and J. G. Fossum, "Viable deep-submicron FD/SOI design for low-voltage applications," Proc. International SOI Conference, pp. 23-24, Nantucket, MA, Oct. 1994

[You88] K. K. Young, and J. A. Burns, "Avalanche-induced drain-source breakdown in silicon-on-insulator n-MOSFET's," IEEE Trans. Electron Devices., vol. 35, pp. 426-431, Apr. 1988

BIOGRAPHICAL SKETCH

Dongwook Suh was born in Seoul, Korea in 1966. He received the B.S. degree from Seoul National University, Seoul, Korea, in 1988, and M.S from the University of Florida, Gainesville, FL, in 1989.

After serving the Republic of Korea Army, he began pursuing the Ph.D. degree in electrical engineering at the University of Florida. He is currently engaged in physical modeling of non-fully depleted (NFD) SOI MOSFET's, and applications in device/circuit design optimization.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Jerry G. Fossum, Chairman
Professor of Electrical and Computer Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Robert M. Fox
Associate Professor of Electrical and Computer Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Mark E. Law
Associate Professor of Electrical and Computer Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Sheng S. Li
Professor of Electrical and Computer Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Chen-Chi Hsu
Professor of Aerospace Engineering, Mechanics and Engineering Science

This dissertation was submitted to the Graduate Faculty of the College of Engineering and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August 1995

Winfred M. Phillips
Dean, College of Engineering

Karen A. Holbrook
Dean, Graduate School